First results from high redshift quasar searches in VIKING

Joseph Findlay

Thesis submitted for the degree of
Doctor of Philosophy (PhD)
of the University of London October 2011

Abstract

This thesis presents the discovery of the first luminous z $\gtrsim$ 6.5 quasars in the VISTA kilo-degree Infrared Galaxy Survey (VIKING). After some basic quality control, quasar selection is investigated via use of initial data supplemented with detailed modelling of the photometric and spatial distributions of stars of spectral type M, L and T, which are known to be the cause of significant contamination in quasar colour selection spaces. Optimised selection constraints are placed on detection significance and morphology and the performance of a traditional colour selection technique is compared to a Bayesian model comparison technique. The latter is found to offer a $\sim$10 per cent gain in completeness over traditional colour selection. Quasar candidates are ranked via Bayesian model comparison and a subset of the highest ranked objects are put forward for follow-up imaging.

In June 2011, 44 high-z quasar candidates underwent deep optical *i*- and *z*- band imaging on the ESO NTT. Just 6 of these candidates were found to have optical colours consistent z $\gtrsim$ 6.5 quasars. Spectroscopic follow-up of these objects is ongoing, but thus far three new quasars have been discovered at redshifts of z=6.5, 6.7, 6.9. This discovery rate is consistent with zero evolution in the rate of decline in quasar space density from z $\gtrsim$ 6.4. This differs from the latest results from UKIDSS. Further results expected from these and other surveys will begin to constrain the true nature of quasar space density evolution in the near future. The discovery of three z $\geq$ 6.5 quasars in VIKING is a significant highlight in the first year of VISTA science operations. These quasars will remain important probes of the high-z universe throughout the next decade.

Contents

List of Figures

List of Tables

To Anthea

Acknowledgements

I would like to extend my deepest gratitude to Will Sutherland for his supervision of this PhD and his contribution to my understanding of many aspects of extra-galactic astronomy involved in this work. Thanks to Jim Emerson for valuable advisory support and to both of the above for the realisation of VISTA.

My gratitude goes to Bram Venemans, who wrote the successful proposal for imaging time on the NTT and secured the follow-up spectroscopy on FORS II, that led to the discovery three quasars.

Staff and students at the School of Mathematical Sciences at QMUL have made the last four years thoroughly enjoyable. In particular I would like to thank Heidar Thrastarson, Moritz Fragner, Adam Christopherson, Roman Pechhacker, Chris Haynes, Chris Watkins, Friedrich Lenz, Simone Fleuren, Stephan Fendyke, Ian Huston, Giammarco Campanella, Oliver Gressel, Enrico Camporeale, Gareth Williams, Yiannis Tsapras, Karim Malik and anyone else I have hopefully not overlooked.

Thanks to my parents Ian and Ruth Findlay for moral support and for allowing me to over stay my welcome in their humble abode and to Katie Evans for forcing me out of bed and into the office on time (plus or minus…) every morning.

Publications

1. J. R. Findlay, W. J. Sutherland, B. P. Venemans et. al., *Selection constraints on high redshift quasar searches in the VISTA kilo-degree infrared galaxy survey*, 2011, MNRAS, in press

2. B. P. Venemans, J. R. Findlay, et. al., *Discovery of three luminous* $z \geq 6.5$ *quasars in the VISTA kilo-degree infrared galaxy survey*, 2011, in preparation

CHAPTER 1

Introduction

1.1 High redshift quasars as cosmological probes

Quasars are the brightest non-transient objects in the Universe and offer powerful probes of cosmology, setting constraints on early structure formation [8, 9, etc.], chemical enrichment [10, and refs. therein], the state of the inter-galactic medium at the end of cosmic reionisation [11], etc. This chapter describes three major research areas to have benefited significantly from discoveries of high-z quasars in recent times. In addition some recent search strategies are summarised and the capabilities of new instruments and surveys relevant to work in this thesis are put into context.

Throughout this work the following values for the Hubble constant and the matter and cosmological constant energy density parameters are adopted; $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.28$ $\Omega_\Lambda = 0.72$, [12].

The passbands of the Visible and Infrared Survey Telescope for Astronomy [VISTA; 13] and and Sloan Digital Sky Survey [SDSS; 14] are referred to in this thesis as u, g, r, i, z, Z, Y, J, H, Ks. Passbands on other systems are differentiated either by the addition of a subscript (e.g. $\mathrm{J_{2MASS}}$) or by explicit statement. In order to avoid confusion with the SDSS z band redshift is denoted non-italic lower case z.

1.1.1 The Hydrogen Atom and Gunn-Peterson Absorption.

Broadly speaking the spectra of quasars are dominated by a striking featureless continuum, which extends over the entire electromagnetic spectrum from radio to the gamma-ray energies. This emission is characterised by the presence of relativistic free electrons [15], which give rise to Synchrotron and inverse Compton emission. The optical/ultra-violet (UV) continuum component is clear from Figure 1.1, which shows the composite rest-frame spectrum of optically selected quasars from the SDSS [1] (dashed and dotted lines show that this continuum has two power law components).

Further evident from Figure 1.1 is that the continuum is adorned with a multitude of

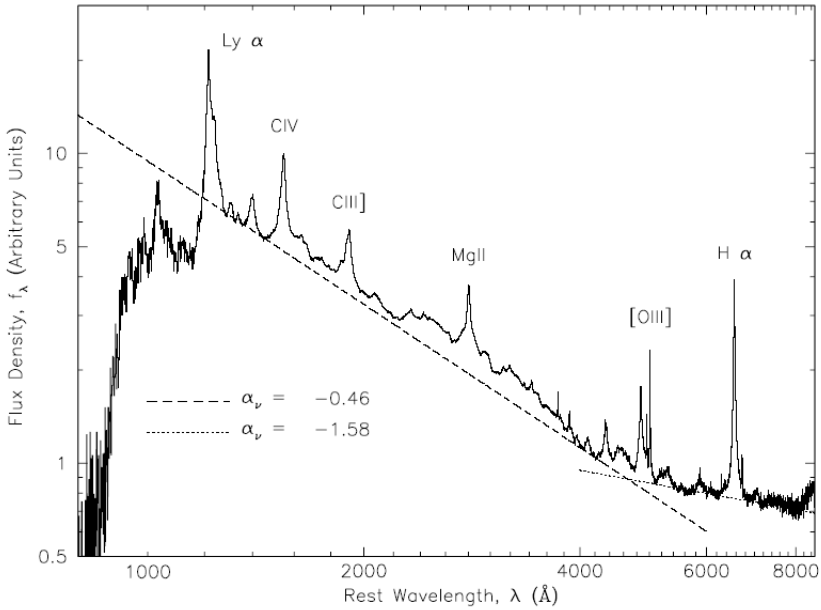


Figure 1.1: A composite spectrum of quasars from the SDSS by Vanden Berk et al. [1]. The spectrum shows the two stand-out features of all quasar spectra displaying a featureless power law continuum adorned with a multitude of spectral emission lines. Quasars are strong emitters of Lyα photons and the continuum extends blueward of the Lyα transition.

spectral transitions, the strongest of which have been labelled. There is a large body of literature covering the origin of these transitions and these discussions are not repeated here [see 16, for a review]. The purpose of presenting Figure 1.1 here is to demonstrate that quasars are strong emitters of Lyman-alpha (Lyα) photons and that the continuum extends blueward of the Lyα transition.

As we shall see, the presence of Lyα in both emission and absorption in quasar spectra has been crucial to their identification in current and past searches. This section briefly reviews the production mechanism of Lyα photons and the signatures of the transition in quasar spectra.

For a first order quantum mechanical treatment, the Hydrogen atom may be modelled as an electron with quantised orbital angular momentum $L = \sqrt{l(l+1)}\hbar$ in a Coulomb potential $V(r) = -e^2/4\pi\epsilon_0 r$, where $l = 0, 1, 2\ldots$ is the azimuthal quantum number, e is the charge on an electron, ϵ_0 is the vacuum permittivity and $\hbar$ is the reduced Planck constant.

The wave equation of the electron contains all the information on the current state of the system and its radial component $u(r)$ is a solution to the radial Schrödinger equation

$$-\frac{\hbar^2}{2m_e}\frac{d^2u}{dr^2} + \left[\frac{l(l+1)\hbar^2}{2m_e r^2} + V(r)\right]u = Eu \tag{1.1.1}$$

where m_e is the electron mass and E is the electron energy.

One can appreciate the qualitative nature of the energy of the system by considering the effective potential term inside the square brackets

$$V_{\text{eff}}(r) = \frac{l(l+1)\hbar^2}{2m_e r^2} - \frac{e^2}{4\pi\epsilon_0 r}, \tag{1.1.2}$$

if $l > 0$, then $V_{\text{eff}}(r)$ is attractive (i.e. directed towards the nucleus) at large r and repulsive at small r. By differentiating equation 1.1.2 one may show that $V_{\text{eff}}(r)$ has a minimum $V_{\text{eff}}(r) = -E_R/l(l+1)$ at $r = l(l+1)r_0$, where $E_R = 13.6\,\text{eV}$ is the ionisation energy often referred to as the Rydberg energy and $r_0 = 0.529 \times 10^{-10}\,\text{m}$ is the Bohr Radius.

In reality of course, the electron may only occupy a discrete set of energy states when in atomic form and transitions between discrete states of differing energy give rise to spectral lines such as Lyα. The wave function of an electron in a bound state is an eigenfunction of equation 1.1.1, which belongs to the energy eigenvalues E. For the interested reader the derivation of the radial Schrödinger equation and the solution of the eigenvalue problem is given a full and clear treatment by Phillips [17]. The results of this treatment show that an electron with angular momentum L in a Coulomb potential has an infinite number of bound states with energies

$$E_{n_r,l} = -\frac{E_R}{(n_r + l + 1)^2} \tag{1.1.3}$$

where $n_r = 0, 1, 2\ldots$ is known as the radial quantum number and by convention bound states have negative energy while ionised states have $E \geq 0$. From equation 1.1.3 it is clear that there exist distinct states with equal values of $n_r + l$ and therefore equal energies. Such states are degenerate and it is normal practise to make the substitution $n = (n_r + l + 1)$, where $n = 1, 2, 3\ldots$ is the principle quantum number. Also apparent from equation 1.1.3 is that as n increases the energy difference between adjacent states decreases and thus transitions between states of high n give rise to spectral lines at longer wavelengths and lower energies. As n approaches ∞ the quantised energy states mimic a continuum. These points are clarified by the energy level diagram for the Hydrogen atom in Figure 1.2.

In quasar spectra the Lyα line exists strongly in both emission and absorption. Lyα is part of a series of transitions known as the Lyman series which arise from transitions between excited states of Hydrogen (with $n > 1$) and the ground state ($n = 1$). The emission lines resulting from the first three of these transitions are Lyα ($n = 2 \rightarrow 1$), Lyβ ($n = 3 \rightarrow 1$) and Lyγ ($n = 4 \rightarrow 1$) and their appearance in quasar spectra is due largely to recombination, whereby an electron in an ionised gas is captured by a proton and cascades down to the ground state emitting a photon of energy $E_{\text{i}1} = E_{\text{i}} - E_1$, i.e. the energy difference between i$^{\text{th}}$ state, where i > 1, and the ground state.

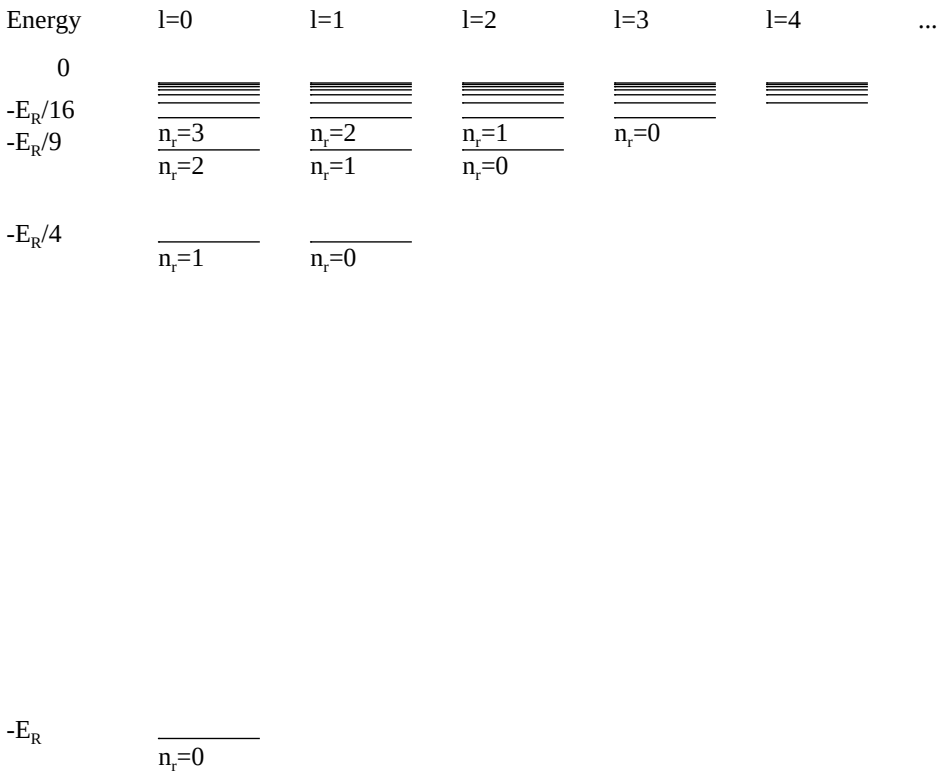


Figure 1.2: Energy levels of the bound states of the Hydrogen atom as given by equation 1.1.3. Each value of the angular momentum quantum number $l = 1, 2, 3 \ldots$ corresponds to an infinite number of states with energies $E_{n_r,l}$ and radial quantum number $n_r = 1, 2, 3 \ldots$

Collisional excitation is also a prominent catalyst of Lyman series emission. In this case an electron is elevated from the ground state energy to a higher energy state via a collisional interaction and a photon of energy E_{i1} is emitted as the electron cascades back down to its ground state.

From equation 1.1.3 the energy of a photon emitted by an electron via a transition from n_i to n_j is $\Delta E = E_R(n_j^{-2} - n_i^{-2})$ giving

$$E_{Ly\alpha} \equiv E_{21} = 10.2\,\text{eV}, \qquad E_{Ly\beta} \equiv E_{31} = 12.1\,\text{eV}, \qquad E_{Ly\gamma} \equiv E_{41} = 12.75\,\text{eV}$$

and since $E = 2\pi\hbar c\,/\lambda$, where c denotes the speed of light and λ denotes the wavelength of the emitted photon, then

$$\lambda_{Ly\alpha} = 1216\,\text{Å}, \qquad \lambda_{Ly\beta} = 1026\,\text{Å}, \qquad \lambda_{Ly\gamma} = 972\,\text{Å}$$

As mentioned earlier, as well as the characteristic Lyman series emission lines in the spectra of quasars, Lyman series absorption features are also apparent and their presence in quasar spectra at high-z are crucial to the quasar search described in this thesis.

Lyman series absorption is brought about by the photoionisation of intergalactic hydrogen by the quasar continuum blueward of the Lyα transition. To appreciate this let us make a rough calculation of the optical depth to Lyα photons along the line of sight in a uniform and homogeneously distributed intergalactic medium (IGM). We start by writing the proper mean number density of hydrogen nuclei at redshift z in terms of the present day baryon density Ω_b;

$$\begin{aligned} \varrho_H(a) &= \frac{\rho_{crit}}{m_H}(1-u)\Omega_b\, a^{-3} \\ \varrho_H &= 1.1\times 10^{-5}\Omega_b\, h^2\,\text{cm}^{-3} \end{aligned} \tag{1.1.4}$$

where u is the fraction of mass in helium, ρ_{crit} is the critical density, m_H is the mass of atomic hydrogen and $a = 1/(1+\text{z})$ is the scale factor.

The integral of ϱ_H over proper length to redshift z gives us the column density Q of atomic hydrogen along the line of sight. In a Friedman-Robertson-Walker Universe the proper line element is given by $dl = c\,da/a\,H(a)$. Therefore

$$Q = c\int_a^{a_0} \varrho_H(a)\,\frac{da}{a\,H(a)} \tag{1.1.5}$$

where $a_0 = 1$ is scale factor at the present day.

The optical depth τ to photons of frequency ν along the line of sight is then the product of Q and the interaction cross-section $\varsigma(\nu)$. Assuming $\varsigma(\nu)$ is highly peaked about the Lyα frequency then $\varsigma(\nu) \approx \varsigma_{\mathrm{Ly}\alpha}\delta(\nu - \nu_{\mathrm{Ly}\alpha})$, where δ is the Dirac delta function. Therefore

$$\tau = c \int_a^{a_0} \varsigma_{\mathrm{Ly}\alpha}\, \delta\Big(\frac{\nu}{a} - \nu_{\mathrm{Ly}\alpha}\Big)\, \varrho_H(a)\, \frac{da}{a\, H(a)} \tag{1.1.6}$$

where $\delta(\nu/a - \nu_{\mathrm{Ly}\alpha})$ accounts for the shift in ν with the expansion of the Universe. For photons emitted at $z \geq 1$ the Universe is matter dominated so the approximation $H(a) \approx H_0\sqrt{\Omega_m}\, a^{-\frac{3}{2}}$ is valid [15] and

$$\begin{aligned} \tau &= \frac{\varsigma_{\mathrm{Ly}\alpha}\, \varrho_H\, c}{H_0\, \sqrt{\Omega_m}}\, \frac{a^{-\frac{3}{2}}}{\nu_{\mathrm{Ly}\alpha}} \\ &\approx 13000 h^{-1}\, \frac{\Omega_b h^2}{0.02} (1+z)^{\frac{3}{2}} \end{aligned} \tag{1.1.7}$$

Where h is the reduced Hubble constant.

At $z \geq 1$ the optical depth of the IGM to Lyα photons would be $\tau \sim 10^5$ and thus for even a tiny content of neutral hydrogen ($< \varrho_{HI}/\varrho_H > \sim 10^{-5}$) one would expect almost complete absorption of the quasar continuum blueward of the Lyα transition. As we shall see this fact is crucial in our ability to identify quasars at high-z with broad band imaging.

In reality transmission of Lyα photons is apparent at some level right up until $z \leq 6$. In fact, approaching this epoch from $z = 0$ the spectra of quasars initially show almost complete transmission. Gradually, approaching higher redshifts transmission becomes patchy and a multiplicity of narrow absorption lines, known as the Lyα forest, is produced as energetic photons are redshifted to the Lyα frequency and are absorbed as they pass through clouds of neutral hydrogen in the *inhomogeneous* IGM. This leads us the conclusion that the Universe has undergone a period of cosmic reionization.

1.1.2 Cosmic reionisation

Throughout cosmic history the ionisation state of pre-galactic hydrogen has undergone two major changes. Initially maintaining an ionised photon coupled plasma, the Universe cooled via expansion, leading to a point at which it is thermodynamically favourable for neutral Hydrogen to exist over plasma. This allowed the photon mean-free path to go to infinity leaving behind the last scattering surface known as the Cosmic microwave background at $z \sim 1090$ (CMB).

Gravitational instabilities grew from primordial density fluctuations, which are present at the surface of last scattering and leave temperature and polarisation anisotropy signals in the CMB. During this process polarisation is generated by Thomson scattering and is seen at small angular scales corresponding to the horizon size at that epoch. The detection of a polarisation signal on scales larger than the horizon at recombination is indicative of secondary processes such as a change in the ionisation state of pre-galactic hydrogen.

The rise of the first population of ionising sources (probably star forming galaxies or possibly active galactic nuclei; AGN) began the process of reionisation. As predicted this process gave rise to a high significance peak in the Wilkinson Microwave Anisotropy Probe's (WMAP) observations of the E-mode polarisation power spectrum at large angular scales [18] with an amplitude proportional to the total optical depth (τ). This is consistent with an optical depth of $\tau = 0.088 \pm 0.015$. Assuming an instantaneous reionisation scenario this would imply a reionisation redshift of $z = 10.5 \pm 1.2$. Reionisation is not expected to have occurred instantaneously, rather the CMB result shows the integrated effects of Thomson scattering and the result is more commonly viewed as the average redshift of reionisation.

The exact nature of reionisation is dependent on the production, distribution and strength of ionising photons and the complex interplay with the topology of the IGM. Theoretical work [e.g. 19–24] has converged on a broad consensus on how reionisation unfolded. The very first ionising sources were initially rare and began to ionise local spheres or "bubbles" of (dense) gas. As more light sources began to populate the Universe, the cosmic expansion caused the IGM density to fall and consequently it became increasingly difficult for ionised hydrogen (HII) to recombine. Ionised regions grew in size and number and eventually they began to overlap causing a sharp increase in the photon mean free path. The final stages of reionisation are expected to have proceeded quickly as ultra-violet photons worked on the final pockets of neutral gas.

The first observational constraints on this process were placed during the last decade via the Gunn-Peterson (GP) test [25], which measures the neutral fraction of IGM along a particular sight line as implied by the absorption of Lyα photons in the spectra of ionising objects at high-z. The test is sensitive to the end of reionisation, since the absorption saturates for neutral fractions of $(\varrho_{HI}/\varrho_H) > 10^{-3}$.

At the time, significant numbers of $z \sim 6$ quasars were being discovered in wide-field optical surveys (section 1.2) and in 2001 the first complete GP trough (i.e. complete absorption shortward of Lyα) was found in the spectrum of a $z = 6.28$ quasar [26]. This was widely interpreted as a clear sign post of the end of reionisation. However, subsequent discoveries soon showed clear transmission in the IGM along sight lines towards higher redshifts [e.g. 27].

The optical depth of the IGM to Lyα and Lyβ photons was measured along 19 sight lines

towards high-z quasars by Fan et al. [11]. They found an abrupt qualitative increase in the rate of evolution of the optical depth of the IGM at $z \simeq 5.5$. An evolution in the scatter of the transmitted fraction with z was interpreted as a sign that they were probing the end of the overlap phase of reionisation. More recently these results have been explained without the need to resort to reionisation overlap [28].

Since the GP absorption rapidly saturates with increasing neutral fraction, further tests of reionisation were developed. In particular the size of dark gaps (i.e. absence of flux) in quasar spectra are explained by variations in the ultra-violet background [29]. As predicted by Paschos and Norman [30] the length of these dark gaps increases significantly at $z > 6$ as the IGM begins to become more neutral.

Further classic tests attempted to measure the extent of the ionised bubbles around quasars [11, 31]. In general it was found that these bubbles increase in size dramatically towards lower redshifts and inferred IGM neutral fractions suggest that reionisation was taking place at these epochs.

It was later shown however that at $z > 6$ resonant absorption by residual H I can lead to an underestimation of the physical size of ionised bubbles which results in an underestimate of the neutral fraction [32].

Whether we are probing the final stages of reionisation with $z \sim 6$ quasars is an open question. Current methodology requires many more independent sight lines and spectra with higher dynamic range [33]. It is now recognised however that quasars as a population cannot give a good average representation of the IGM during reionisation since they reside in regions that contain an over-density of ionising sources [34]. This should not detract from the use of quasars as probes of these interesting regions but it looks likely that reionisation as a global phenomena will be better probed in future by gamma-ray-bursts [35] and deep galaxy surveys [36] afforded by the new near-infrared capabilities of the Hubble space telescope (HST).

That said, there is still an important role for high-z quasars as probes of reionisation in the most biased regions of the Universe. This was illustrated in no uncertain terms with the detection of a Lyα damping wing in the spectrum of the highest redshift quasar yet discovered [37]. Furthermore, the scale of the ionisation zone around this quasar shows clear departures with those of $z \sim 6$ quasars. Further examples of $z \sim 7$ quasars are required to determine if these results are mirrored elsewhere.

The introduction of low frequency radio arrays will revolutionise current knowledge in the area through redshifted 21-cm observation, allowing full 3-dimensional imaging of the IGM on all scales (linear and non-linear). The imaging of the neutral IGM will allow us to distinguish between competing reionisation models. There is a large body of literature describing the different projects and tests on the IGM with 21-cm observations, much of this is neatly summarised by Morales and Wyithe [38].

Among the most easily detected and interpreted signatures will come from the distinction between redshifted 21-cm signals from large H II bubbles driven by high-z quasars and the surrounding IGM [39]. Quasars will become valuable probes of the IGM during the end of the overlap period and quasars discovered in current wide field imaging surveys thus represent ideal targets for searches in forthcoming 21-cm surveys. Coupled with this, rest-frame ultra-violet observations of high-z quasars will continue to provide good qualitative constraints on the end of the reionisation epoch.

1.1.3 Black hole growth and evolution

In 1978 a review paper by Rees [40] considered three different scenarios for the seeding and subsequent pathways to growth for supermassive black holes (SMBH); growth from stellar mass black holes, direct collapse of gas in the centres of protogalaxies and runaway collapse of dense star clusters. With the more recent addition of the hypothetical primordial black holes produced by density fluctuations during the Big Bang, all continue to be regarded as viable candidates for the progenitors of SMBH.

The vast majority of work in this field is still theoretical and there are a number of detailed reviews covering the theoretical intricacies [e.g. 41]. The single most compelling observational constraint on this field is the observation of quasars powered by $10^9\,M_\odot$ black holes (BH) just a Gyr after the Big Bang.

A difficulty arises from classical argument; the theoretical limit on the mass accretion rate of a gravitationally accrediting body emitting radiation isotropically comes about through the balancing of radiation pressure with gravity via the Eddington limit. The implied accretion rate has a mass doubling time scale of 45 Myr, the Salpeter time.

The recent important discovery of a $z \simeq 7$ quasar harbouring a $2 \times 10^9\,M_\odot$ SMBH [37] is significant because within the confines of concordance cosmology, there is little more than 0.7 Gyr from which to form such an object. If one assumes that the SMBH grew from a seed of $10^2\,M_\odot$ (a population III star remnant for instance), then the mass of the object at $z \simeq 7$ requires 17 e-folding times, which is pushing the limits of the Salpeter argument. Therefore the discovery of quasars at high redshifts places strong constraints on mechanisms for massive BH seeding and growth.

Observationally, further constraints on these processes come principally through the quasar luminosity and mass functions. The quasar luminosity function (QLF) charts the space density of luminous quasars per magnitude interval in redshift space. Therefore the QLF contains information regarding the rate of mass build up of SMBHs and the reproduction of the QLF is seen as a minimum requirement for acceptable cosmological simulations [e.g. 42].

At $z \sim 6$ most SMBHs are in an active phase of accretion and most moderate to high mass SMBHs are accreting close to the Eddington limit [43]. This allows the inversion

of the QLF to obtain the SMBH mass function (MF).

The latest determination of the MF [43] has placed the latest observational constraints on the build up of massive BHs in the early universe. Scaling relations between the mass of SMBHs and their host galaxy properties in the local Universe [44] are interpreted as a sign that both grow synchronously, subject to the same physical processes during their formation. It is surprising to find from the latest MF, that the SMBH mass density appears to evolve a factor 100 times more than the stellar mass density between $z = 0$-6.

It has been suggested that most e-folds in mass are acquired through accretion in the cores of dense dusty proto-galaxies before the quasar luminous phase. This would account for the vast discrepancies in quasar and stellar mass build up at early times. This is also consistent with one basic evolutionary paradigm in which SMBHs begin life embedded in the gas rich progenitors of local 'red and dead' ellipticals. Massive rates of star formation occur co-eval with BH growth obscured from view until a combination of feedback, star-formation and BH feeding allows the breakout of a visibly luminous nucleus and the beginning of the quasar phase.

Such systems would perhaps be detectable in hard X-rays, but despite major efforts, there is still no evidence of the existence of these heavily obscured quasars at high redshift, even in deep stacking analyses [45].

If they exist it will be important to try and observe these systems directly, since during luminous accretion SMBHs are acquiring most of their present day mass, erasing valuable information regarding their initial conditions.

Volonteri [46] discusses Monte-Carlo simulations of SMBH growth in a merger driven accretion scenario via a Press-Schechter formalism. Starting with different initial seeds each object accretes an amount scaled to the mass-dispersion velocity relation of its host. Expectations for the integrated co-moving mass densities in SMBHs seeded by early (i.e. population III stars; primordial BH were not considered) and late time objects (gas and stellar cluster collapse), are indistinguishable up to the peak of quasar activity at $z \sim 2.5$, but become rapidly more distinguishable with increasing look-back time. The related observable; the accreted mass density, does not diverge until look-back times corresponding to $z \gtrsim 10$.

One must determine the accreted mass density in SMBHs from observations of both obscured and luminously accreting SMBHs. We are already able to detect the most luminous quasars but the steep faint end slope of the quasar luminosity function would imply that these are only the tip of the iceberg. In the future the James Webb Space Telescope (JWST), will detect the more typical quasars with masses in the range 10^5 - $10^6\,M_{\odot}$. The Atacama Large Millimetre Array (ALMA) will peer through the dust to perhaps detect thermal emission from future quasars in obscured star forming galaxies. Dark energy orientated projects such as the Large Synoptic Survey Telescope (LSST)

and the European Space Agency's Euclid mission are expected to detect significant numbers of luminous quasars at $z \sim 10$ on a similar time frame.

Thus in the not too distant future we can expect this field to benefit from an observational complement to theory, with quasar research playing an "active" and prominent role.

1.1.4 Massive galaxy formation and assembly

Correlations between SMBH masses and the properties of their host galaxies are well established [44, for a review]. Coupled with the realisation that SMBHs reside at the centres of most massive galaxies [47], this points to a scenario in which the evolution of the nucleus and spheroid are governed by the same physical processes. Indeed galaxy mergers and interactions are believed to be a catalyst for both BH growth and star formation, providing the dense environments favoured by stellar collapse as well as transporting fuel to the galactic nucleus to trigger the quasar luminous phase.

The peak in quasar activity occurred at $z \sim 2.5$ and actively accreting BHs exist at $z > 6.5$. Around the same time massive galaxies were rapidly forming stars and their remnants are visible today as old population gas poor ellipticals. Their gas rich progenitors, formed at high redshift, are commonly thought to be ultra-luminous-infrared-galaxies (ULIRG). The initial evolution of these galaxies is enshrouded in dust in which there is evidence for massive rates of star formation and Compton thick X-ray sources [48]. This allows the entire system to undergo rapid chemical evolution in relative obscurity, before a combination of star-formation, BH feeding and feedback from the quasar outflows allows the breakout of the visibly luminous nucleus and the beginning of the quasar phase.

The study of quasar host galaxies is offering support for this evolutionary paradigm. For instance elemental abundances in broad lines originating from the nuclear region (Broad Line Region; BLR) are a direct probe of past star formation during the non-luminous phase of evolution.

Among the important results is that there is a lack of metal evolution in the rest-frame ultra-violet spectra of quasars, which show metallicities typically 2-5 times solar at all redshifts [37, 49, 50]. Seemingly all SMBH hosts have reached their heights in chemical enrichment prior to becoming luminous quasars, supporting the case for a dust obscured phase of evolution.

Secondly a correlation between metallicity and luminosity is understood as a manifestation of the well known mass-metalicity relation in normal galaxies [51], provided of course that the galaxy-BH mass correlation holds true at high redshift. This would suggest that the most massive quasar hosts build up their stellar mass more efficiently than their low mass analogues.

The use of quasars as probes of these processes is unique in this respect since they offer the only chance to obtain non-transient high signal-to-noise spectroscopy. Normal field galaxies are, for now, just too faint to attempt similar observations [52], while GRBs are short lived.

Ongoing star formation in quasar hosts can be probed via radio observations of the cool dust from which stars are born. Similarly warm dust resulting from the heating generated by star formation emits in the sub-millimetre. In general these observations imply the presence of huge dust reservoirs $\sim 10^8\,M_\odot$ and star formation rates of $10^3\,M_\odot year^{-1}$ [53] implying huge star burst activity co-eval with BH growth. High-z quasar hosts show far-infrared excess compared to their low-z counterparts and their spectral energy distributions (SED) are consistent with the far-infrared to radio correlations found in low-z star forming galaxies. The implied high rates of star formation in these galaxies are consistent with the picture of quasars in their luminous phases having just broken through a vale of dusty star formation.

Regions of highly excited CO show evidence of warm gas on scales of kpcs with densities similar to that seen in Galactic giant molecular clouds [54]. The findings are consistent with the picture from the BLR over galactic scales suggesting that even the host galaxies of quasars are highly enriched at early times.

These are but a few of the recent radio and far-infrared studies of the oldest galaxies which have focused their attention almost exclusively on quasar hosts. With the exception of a small number of gamma-ray bursts, quasars are the only high-z objects with spectroscopic redshift determinations that allow for bandwidth limits to be placed on millimetre line observations. Radio observations are also limited to the most massive galaxies at high-z so quasar hosts are natural targets since they represent the most biased events of galaxy formation in the Universe.

Huge dust reservoirs present in galaxies just 1 Gyr after the Big Bang must have been built up rapidly. The presence of hot dust in quasar hosts is indicated by their near-infrared continua excess, a result of direct heating by the central engine, which is ubiquitous in quasar hosts at lower redshifts. Tentative recent results have found systems deficient in hot dust in quasar hosts at high-z [55]. At low-z dust in quasar hosts has been reprocessed and recycled between galaxies and the intergalactic medium and thus BH-mass is not a good diagnostic of dust mass, however at high-z there is a clear trend towards the dustier systems harbouring more massive BHs again pointing to co-eval growth. Dust deficient systems harbour the least massive BHs but have high Eddington ratios, suggesting that they are undergoing rapid accretion but are too young to have built detectable levels of dust in the vicinity of the central engine.

These results are so far limited to a few examples at high-z and further more, examples of (all be it less extreme) hot-dust deficient quasars have been found at $z < 4$ [56]. Other mechanisms can be invoked to explain the presence of these systems at significantly

lower redshifts, but if there is truly an evolutionary link one would expect to find that the frequency of these objects increases at higher redshifts.

The use of hot dust as a probe of the interplay between galaxy assembly and growing BHs is yet another key diagnostic unique to quasar hosts.

1.2 Optical & near-infrared quasar searches

The first quasars were discovered in the third Cambridge catalogue of bright radio sources by virtue of their near spatial coincidence to star-like objects [57]. Soon after the initial discoveries Ryle and Sandage [58] identified a number of sources with quasar like properties lacking radio emission. It soon became clear that the radio loud quasars represent only a $\sim$10 per-cent subset of the entire class. Despite this, significant numbers of radio selected quasars exist in the literature, owing to the availability of deep radio surveys that cover extensive areas of sky [e.g. 59].

As the need for equally large radio quiet data sets developed, attention was paid to optical quasar selection. The field benefited greatly from the arrival of the automatic plate measuring machines and multi-fibre spectroscopic facilities. Early techniques relied on photometric detections of the ultraviolet excess (UVX) making quasar U-B colours distinct from stars [e.g. 60, 61]. Unlike radio searches, UVX selection is subject to sizable levels of contamination, principally from white dwarfs and other hot stars. Consequently, spectroscopic conformation is necessary to identify false positives. Nevertheless the UVX technique is still reasonably efficient (especially at faint magnitudes where the number of contaminants decreases appreciably) and has remained effective in the identification of quasars up to $z \lesssim 2$. Beyond $z = 2$ the Lyα transition is shifted out of the ultra-violet and the stellar and quasar U-B colours become indistinguishable.

Further progress was made most notably by slitless spectroscopic and multi-colour surveys. Slitless spectroscopy identifies quasars via their strong emission lines on low-dispersion objective-prism photographs. The technique was transferred to large four meter class telescopes by use of either a grating prism combination (grism) or a grating-lens combination (grens). The introduction of these elements significantly improved the sensitivity of slitless spectroscopy. The approach has delivered large data sets of objects spanning ranges of approximately 0 to 4.5 in redshift [e.g. 62, 63].

The use of slitless spectroscopy has become somewhat less popular in recent years due to the fact that it is plagued by unwanted selection effects. The technique is naturally affected by the strength of emission lines, which introduces redshift related bias as strong lines are shifted in and out of the relevant passbands. Line detection is also highly dependant on seeing conditions, detector response and wavelength dependent dispersion. Further difficulties arise in placing a limiting magnitude on surveys. The conclusion is that most slitless spectroscopic surveys are rather incomplete.

The vast majority of confirmed quasars have been selected by virtue of their location in multicolour space. The technique was pioneered in the late 1980s and early 90s [64–66] and led to the discovery of the first z$\sim$4 quasars [67], paving the way for truly massive multi-colour surveys of modern times.

In the last decade, optical surveys such as SDSS and the Canada France High Redshift Quasar Survey [CFHQS; 68] have discovered ~ 50 quasars at $z \sim 6$. At the single scan sensitivity limit of the Sloan camera ($z \simeq 21.0$) the surface density of $z \geq 6$ quasars brighter than $z \leq 21$ is of order $\sim 0.01\,\mathrm{deg}^{-2}$, so this represents a considerable achievement.

The key to the photometric detection of a strong $z \sim 6$ quasar candidate is to isolate the characteristic flux break blueward of the Lyα transition brought about by the thickening of the Lyα forest towards higher redshifts. Colour selection techniques exploit this break by placing a blue observing band across the Lyα transition and a further band just redward. Objects suitable for follow-up spectroscopy should then present themselves as extremely red faint point sources. A carefully placed colour cut is usually sufficient to confine the most interesting sources to a region of colour space distinct from most main sequence stars and galaxies. In principle, this facilitates candidate selection in just two passbands, but in practise a third passband is necessary to permit a measurement of the continuum level redward of Lyα as a means of reducing contamination from cool degenerate stars. This is the technique pioneered by the SDSS and adopted in the CFHQS; both make initial selections in $i - z$ and take followup J-band imaging of promising candidates prior to spectroscopic confirmation.

At $z > 6$ Lyα begins to shift out of the Sloan z band and while some current projects like the Panoramic Survey Telescope and Rapid Response System [Pan-STARRS; 69] aim to progress towards higher redshifts in a predominantly optical parameter space (by selecting i dropouts) faint optical detections make it difficult to reject the numerous Galactic stars with scattered quasar like colours that can out number high-z quasars by a factor of 10^4.

To this end a number of other surveys have taken a different approach, employing the near-infrared as their selection space, a tactic which has had recent encouraging success in the UKIRT Infrared Deep Sky Survey [UKIDSS; 70]. One of the innovations behind the UKIDSS photometric system, was to recognise that near-infrared filter combinations incorporating zJH and K alone are inadequate for $z > 6$ quasar selection, since cool star spectra tend to peak in the near-infrared and crowd the quasar colour-colour locus [71]. The solution was to introduce an observing band intermediate between z and J and was put forward by Hewett et al. [72]. The Y band filter was optimised for this purpose and has been used to great effect in the UKIDSS Large Area Survey (LAS), where seven $z > 6$ quasars have been identified up to the UKIDSS data release eight [37, 73–76].

Figure 1.3: VISTA, overlooked by the VLT telescopes. ESO Cerro Paranal observatory, Chile (credit: G. Hüdepohl).

Working in the same parameter space as UKIDSS is the European Southern Observatory's (ESO) Visible and Infrared Survey Telescope for Astronomy [VISTA; 13], a 4-m class wide field survey telescope located in Chile. More than 80 per cent of VISTA time is devoted to running a suite of six public surveys, which delivered their first public data releases to ESO in mid 2011. Several of these projects are expected to return a significant number of high-z quasars. This thesis proposes to search VISTA data for the most distant quasars yet discovered.

1.3 VISTA

At the time of writing, VISTA is the latest addition to ESO's fleet of world class telescopes. VISTA is currently (in 2011) the world leading facility for wide-field large area surveys in the near-infrared.

VISTA was originally conceived and developed by a consortium of 18 UK universities and was transferred to ESO as part of the UK's accession agreement.

The VISTA site is located on a peak at ESO's Cerro Paranal observatory in Chile some 1500 m from the main summit on which is housed the ESO Very Large Telescopes (VLT, figure 1.3). The area of the Atacama desert host to VISTA is believed to be the driest place on Earth and as such is an ideal site from which to base a near-infrared telescope. Lying some 12 km inland from the Pacific coast and 2,500 m above see level the site is extremely remote. Excellent atmospheric conditions ensure that the sky is photometric 78 per cent of the year round and that the seeing full-width-half-maximum (FWHM) is 0.66 arcsec at the 50th percentile. Here the working design of VISTA is described, the section broadly summarises an in depth and detailed description given by Emerson et al. [13] and references therein.

The design of the telescope and camera were optimised jointly for efficiency, image

quality and sensitivity to fully exploit the followup capabilities of current 8-m class telescopes. A large field of view is therefore a necessity and VISTA was designed as a 4-m class telescope. The overall solution is a quasi-Ritchey-Chretien arrangement with the joint design of both telescope and camera. There is significant aberration from the two mirrors which is corrected in the camera by three field corrector lenses. Thus the telescope is not conducive to the mounting of more general instruments, but in tandem with the VISTA infrared camera (VIRCAM) the overall system introduces negligible chromatic aberration over a 1.65 $\deg^2$ field of view.

VIRCAM incorporates a novel design feature to reject infrared background. Conventionally this is achieved via a cold stop design, but because of design issues with the cold stop route – the use of lenses was rejected due to manufacturing difficulties, while mirror based designs lead to obscuration and accommodation issues – instead VIRCAM incorporates a long cryostat baffle enclosure at the front of the camera. The design ensures that the solid angle of warm exterior seen by the detector array is as small as possible and does not incorporate any part of the telescope other than the primary mirror. A warm reflective baffle is incorporated around the secondary to block sky-emission. The result is that the detectors see only the thermal contributions from the primary, secondary, spider and cryostat window, which would be unavoidable in a cold stop design, plus a small excess from the reflective warm baffle. A cutaway of VIRCAM showing the baffle arrangement is given in the left panel of figure 1.4

The detector array contains 4×4 Raytheon VIRGO 2048×2048 HgCdTe infrared detectors, the largest infrared array manufactured to date. The operating wavelength range is from 800 to 2500 Å. The detector quantum efficiencies (QE) are high and fairly flat with wavelength; there is a gradual decline from $\lambda \sim 1.2\,\mu$m to $0.9\,\mu$m, followed by a sharp cutoff at $\lambda \sim 0.81\,\mu$m due to absorption in the CdZnTe "substrate" (see Figure 1.5). With the 4-m f/1 primary combined with the 1.24-m secondary the resultant beam size is f/3.26, which corresponds to 0.34 arcsec on-sky pixels of $20\,\mu$m in scale, this makes for excellent image quality at the sites 0.66 arcsec median seeing.

Since the array is not edge buttable, there are significant gaps in the focal plane. A single integration therefore produces a sparsely sampled image of the sky referred to as a 'paw-print'. The space between detectors is optimally filled in via a minimum of six sequential pointings. The optimal dithered offsets are; starting with an exposure at the initial position, two dithers in the y (azimuthal) direction a distance 0.475 of a detector height, with an exposure at each offset position, and then a dither offset in the x (equatorial) direction a distance 0.95 of a detector width, with a fourth exposure. Finally, there are two more dithered offsets and exposures in the $-y$ direction each offset from the last a distance 0.475 of a detector height. The right hand panel of figure 1.4 clarifies this procedure.

The filters are placed across the detectors via 1.37-m diameter filter wheel. The phys-

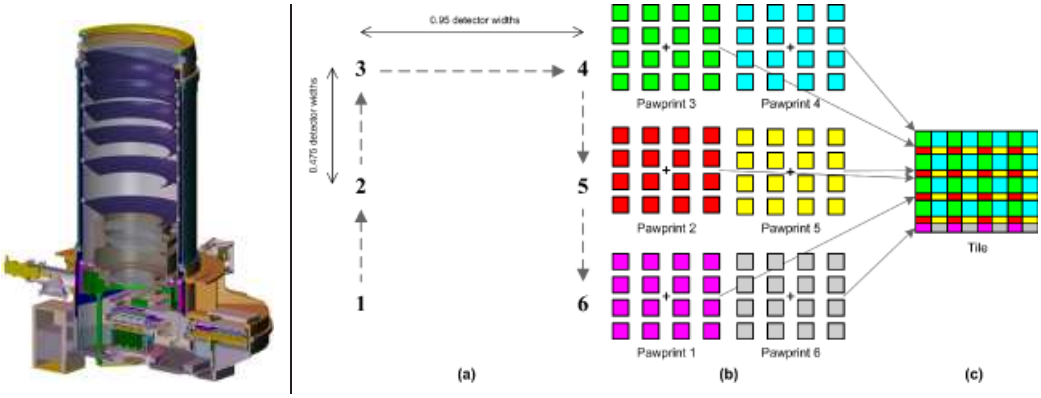


Figure 1.4: Left: VIRCAM cutaway showing the cryostat baffle arrangement and corrector lenses (credit: VISTA). Light passes through the camera window at the top of the baffle tube, through the baffle tube, through the three corrector lenses and is focused on to the detector having passed through the filters. Right: The full tile image (c) is made from a series of six paw prints (b) by offsetting each pointing (a) such that each sky position is covered at least twice on two different pixels (credit: VISTA).

ical filters for VISTA each comprise 16 separate pieces of glass (one per detector) in a metal framework, since monolith filters would need to be over 300 mm corner-to-corner which could not be manufactured at the time of ordering (and, probably not today). The current filter setup is defined for the ESO VISTA public surveys and incorporates five broad band filters at Z, Y, J, H, Ks and a narrow band filter at 1180 Å.

In comparison with the UKIDSS Wide Field Camera (WFCAM), the Z, Y and H filters are very similar; the Ks filter is preferred over the WFCAM K filter due to the warmer site at Paranal compared to Mauna Kea. The VIRCAM J filter has very similar central wavelength but is slightly wider than WFCAM J.

From the above, there are no sharp features in the telescope and instrument responses in the bandpasses in use, i.e. the detector blue cutoff is shortward of the blue edge of the Z filter, and the detector cutoff at $\sim 2.5\,\mu$m is longward of the Ks filter, so the effective bandpasses are primarily defined by the filter transmission curves, with small modifications due to the coating ripple and some minor atmospheric absorption features inside the filter passbands.

Each component of the system throughput; that is the telescope, detector QE, mirror reflectivity and atmospheric transmission, as well as the combined system throughput is shown in figure 1.5.

1.4 VIKING

VISTA is currently running a suite of six public surveys, which deliver public data releases to ESO. Table 1.4 summarises each of the surveys by area, nominal depth and filter coverage and figure 1.6 shows the survey footprints overlaid on a Two-micron all

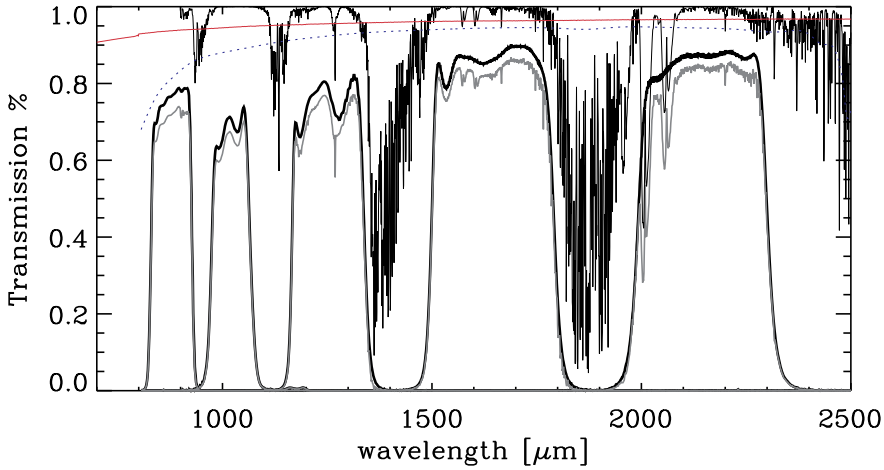


Figure 1.5: Components of VISTA throughput; atmosphere (thin black), primary and secondary mirror reflectivity (red), detector quantum efficiency (blue), filters (thick black). The component transmission curves are available from ESO and were combined to give the effective throughput (grey).

Survey	Area (deg^2)	Depth (5σ Vega)
UltraVISTA	0.73	Y=26.1, J=25.7, H=24.7, Ks=23.8
VIDEO	12	Z=25.2, Y=24.0, J=23.6, H=22.6, Ks=21.7
VMC	184	Y=21.9, J=21.4, Ks=20.3
VVV	520	Z=22.7, Y=22.0, J=21.0, H=19.0, Ks=18.9
VIKING	1500	Z=22.6, Y=21.7, J=21.1, H=20.1, Ks=19.4
VHS	20000	Y=20.6, J=20.3, H=19.2, Ks=18.2

Table 1.1: A summary of the VISTA public surveys. The quoted depths are theoretical, the measured depths vary from these especially in the J-band, see the end of the current section for realistic VIKING depths.

sky survey [2MASS; 77] image of the sky. A full description of each of the surveys and their science goals is given in [78].

The more relevant surveys for the detection of z $\gtrsim$ 6.5 quasars are the VISTA deep extragalactic observations survey (VIDEO), the VISTA hemisphere survey (VHS) and VIKING. VIDEO and VHS are more sensitive to the faintest and brightest quasars respectively. The brightest quasars are intrinsically rare and they are hosted by the most massive galaxies in the universe. These galaxies represent 5 - 6σ peaks in the cosmic density field and in order to discover significant numbers, a survey needs to be extremely wide but may be relatively shallow. The VHS fulfils these demands and is expected to contain several bright z $\gtrsim$ 6.5 quasars.

VIDEO will be sensitive to quasars hosted by less extreme systems, such systems are ubiquitous throughout the early Universe, but until now have been invisible to our telescopes since they are simply extremely faint. VIDEO is expected to contain perhaps one or two of these objects.

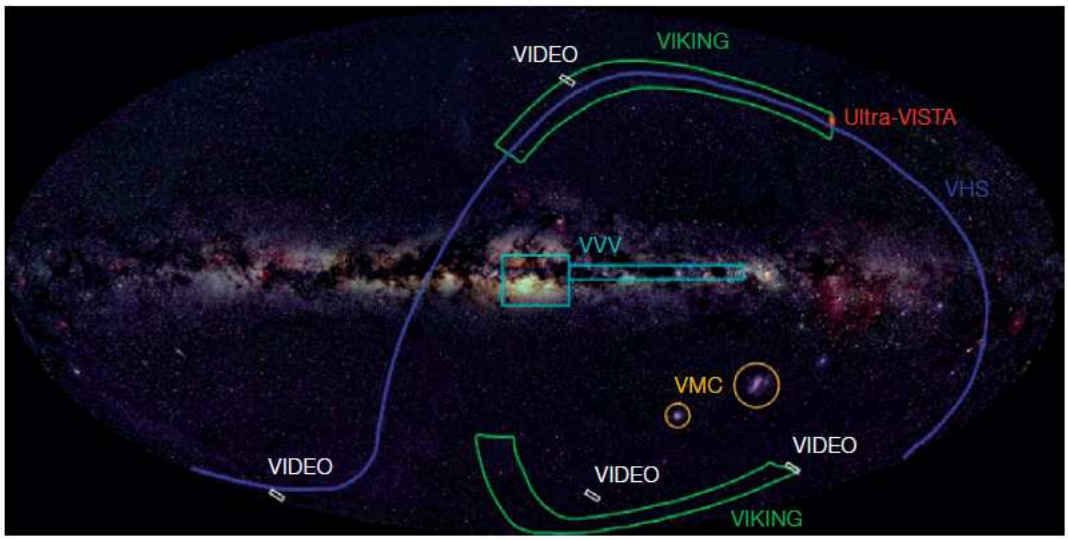


Figure 1.6: Footprints of the VISTA public surveys overlaid on an all-sky image from the 2MASS all-sky survey (credit: ESO).

VIKING forms a natural intermediate between VHS and VIDEO by targeting patches totalling 1500 deg^2 with coverage mostly in the southern hemisphere. VIKING will look for bright quasars at depths of $\sim$ 1 mag fainter than the VHS and is expected to return a significant ($>$ 10) number of high-z quasars. The exact sky coverage is for the most part rigid but there is some flexibility in allowing for maximum synergy with other projects as they and VIKING evolve.

At the time of writing the planned sky coverage comprises two main stripes; an equatorial stripe along the north galactic cap (ngc) at 10 h $<$ Ra $<$ 15 h30 m, $-$05 d $<$ dec $<$ $+$04 d, and a southern stripe crossing the south galactic pole (sgp) at 22 h $<$ Ra $<$ 3 h30 m, $-$36 d $<$ dec $<$ $-$26 d; plus a smaller outlier field incorporating the Galaxy and Mass Assembly [GAMA; 79] 09 field.

This sky coverage includes almost all of the Two-degree Field Galaxy Redshift Survey [2dFGRS; see e.g. 80], and is common with that of the planned VLT survey telescope Kilo-degree Survey [VST-KIDS; 78], an optical imaging survey over the Sloan u, g, r and i passbands. Upon completion, VIKING and KIDS will yield a nine band optical to near-infrared imaging survey with significant improvements in resolution, and gains in depth of $\sim$1.4 mag and $\sim$2 mag over UKIDSS-LAS and SDSS respectively. The VIKING/KIDS areas also contain all of the southern and equatorial Herschel-ATLAS fields [81], and are optimally placed for followup from southern telescopes including ESO telescopes such as the VLT, ALMA and the Extremely Large Telescope (E-ELT).

To achieve VIKING's science goals (which other than high-z quasars include accurate photometric redshifts, weak lensing analysis and the observation of baryon acoustic oscillations, galaxy cluster searches, the study of galaxy stellar mass content and searches for cool stars) , coverage in all five of VISTA's broad band filters is required. The best

seeing at Paranal is not required as morphological analysis of extended sources for galaxy structure studies will come from the VST's *r*-band, however the seeing must not degrade colours measured in the default 2-arcsec aperture and the requested seeing threshold is $\leq$ 1.0 arcsec. This is sufficient to allow most faint point sources, i.e. stars and importantly quasars, to be clearly distinguished from galaxies by their morphology given a detection in multiple VIKING bands alone.

The observation strategy is mostly dictated by the fact that the VIKING Z-band observations, must not lag the KIDS coverage, which is the faster of the two surveys. Thus observation in all five passbands during the same observing block is undesirable and instead each pointing is observed in two visits, whereby the J-band observations are split into equal time slots such that Z, Y and J1 are obtained first and J2, H, Ks are observed a few weeks or months later.

This arrangement is beneficial to the quasar search since high-z quasars are faint in Z. Proper motion objects move between observations separated by long time lags and this can cause them to look like Z band dropouts (suggesting a high redshift) where an object has moved between Z and J observations. The fact that Z and J are taken near simultaneously avoids any chance of spurious Z dropouts being mistaken for high-z quasar candidates. The second pass in J a few weeks later will identify objects with small proper motions, which can pass in front of stars, altering their measured photometry.

VIKING will detect in excess of 6×10^6 stars and 20×10^6 galaxies. The survey is currently delivering 5σ limiting depths of $Z \simeq 22.4, Y \simeq 21.4, J \simeq 20.9, H \simeq 19.9$ and $Ks \simeq 19.3$ (Vega); slightly below nominal. Nevertheless this architecture forms an ideal working parameter space for the discovery of a significant sample of $6.5 \leq z \leq 7.5$ quasars. This thesis concentrates on the search for high-z quasars in VIKING.

CHAPTER 2

Data quality control

The first full depth VIKING catalogue became available to the VIKING working group in April 2011. Users of this initial release were in effect trialing previously untested products so before attempting any science, it was a good idea to obtain a broad overview of the integrity of the pipeline processed photometry and identify any quality issues in the catalogues that may inhibit quasar searches. The distribution of catalogue data to the community is through the VISTA Science Archive [VSA; 82] and this is the starting point for the analysis of all data in this thesis. Thus before analysis of any data for quality issues or otherwise it is useful to look briefly at the VSA.

2.1 The VISTA science archive

Over their lifetimes the total raw data collected by the six VISTA public surveys will amount to several hundred terabytes. For the typical computational resources of the average astronomer, bulk download and processing on this scale is simply impossible. For this reason the VSA was developed as the primary database of all image and catalogue data products generated by VIRCAM. The VSA is a relational database management system, matching data based on common characteristics and presenting it grouped on these characteristics in a related set of tables. Users may query and download data remotely via the provision of custom built tools that allow one to define data subsets and optionally to process query results on the fly.

The VSA is one component of the VISTA Data Flow System [VDFS; 83], which also comprises of the VISTA [84] pipeline at the Cambridge Astronomical Survey Unit (CASU). The VSA inherits pipeline processed data in the form of reduced images and calibration frames, and information collectively known as 'meta-data' encompassing e.g. the observation dates, the environmental and weather conditions, the instrumental performance at the time of each observation, etc.

As discussed in chapter 1, VIRCAM covers a 1.65 degree diameter field of view with sixteen non-buttable 2048 $\times$ 2048 pixel infrared detectors. Hence a minimum of 6 se-

quential offset pointings are required to give a satisfactory sky coverage. The primary science image component in the VSA is thus a tile, produced by stacking the individual dithered pointings consisting of the sixteen detector images collectively known as a Multiframe to fill in the gaps between the individual detectors. In the case of VIKING, tiles can be simply stacks from a single observation block or deep stacks where the same pointing has been visited at two or more distinct epochs, as is the standard for all VIKING J-band observations.

Each tile image is supplied with single passband extracted catalogues which contain the basic photometric measurements and measurement errors, along with 80 or so other standard quantitative attributes describing various aspects of e.g. photometry, morphology, astrometry and quality control fields such as atmospheric seeing and observational zero-points.

The full depth catalogues are extracted from these tiles and subsequently band merged whereby passband level or single epoch detections are combined to produce multi-colour, multi-epoch records. A subset of the most useful standard attributes are propagated into merged catalogues and where appropriate these attributes are recalculated to reflect the extra information provided by the inter-detector, -band and -epoch associations. The result of these curation steps is a set of linked tables containing the data and images relevant to each imaging event, preliminary processing task, individual detection to full band-merged source.

The VSA may be queried via the Structured Query Language (SQL) and several other interactive web forms allowing a high degree of versatility.

2.2 Repeatability

With a VIRCAM integration producing a sparsely sampled paw-print image of the sky, 6 offset pointings are required to fill the gaps between detectors. The 6 offset steps are combined resulting in a tiled image of the sky covered by a minimum of two pixels. The same approach is followed on subsequent nights so that one may check the reproducibility of the photometry by comparing measurements of the same sources between both paw-prints and tiles.

Figure 2.1 shows the root-mean-square (rms) residuals of repeat observations in magnitude bins of width 0.5 mag. Since the mean is more sensitive to outliers, the rms is calculated here as rms $= \sigma_m/\sqrt{2}$, where σ_m is the standard deviation of residuals in magnitude bins and is calculated via the median absolute deviation (MAD) as $1.48 \times$ MAD.

Objects within 10 arcsec of detector edges are omitted from the selection. The VIKING survey requirement is for a 2 per cent uncertainty on the photometry of bright stars

taken on photometric nights, which corresponds to an rms residual of ~ 0.02 in magnitude and is shown by each of the blue lines in figure 2.1. Crosses are the rms residuals while red-points are the estimated errors from the archive which are based on source counts and local background and do not contain contributions form the calibration procedure. Therefore they under-predict the rms residuals, especially at the bright end.

The overall indication from figure 2.1 is that the nominal high quality calibration requirements are broadly speaking, achieved in the initial VIKING release and one can be confident of scientific inferences based on VSA photometry.

2.3 Offsets from the Vega system

VISTA photometry is calibrated by CASU with reference to a set of 2MASS standards, which have been measured with reference to the spectrophotometric standard star Vega ($\alpha-$Lyrae). This implies that Vega is a zeroth magnitude star in all VISTA bandpasses. The CASU calibration involves measuring the offsets between the 2MASS standards and VISTA observed stars via a set of colour equations[1]. Since 2MASS photometry was measured in the J, H, Ks bands alone, the basic assumption is that the 2MASS colour equations can be linearly extrapolated to cover the VISTA Z and Y bandpasses. Any divergence from linearity will result in an offset from the Vega system in the extrapolated wavelength region.

A similar occurrence was noted by Hodgkin et al. [85] in UKIDSS data release (DR) 2, where a rigorous analysis of the Z, Y, J, H and K photometry against the overlapping SDSS footprint found a significant offset in the Y-band. A similar analysis for VISTA will no doubt be attempted when the catalogues are diverse enough to make precise measurements. However in later chapters of this thesis, we shall be synthesising Vega zeroed photometry and it is important to at least tentatively determine any offsets present before we begin.

At the time of this analysis the VSA contains the latest v1.0 VIKING catalogue, containing data up to July 2010. The release comprises of some 200 $\deg^2$ of imaging in Z, Y, J and 120 $\deg^2$ of imaging in all five bands; 90 $\deg^2$ of this has complementary SDSS overlap. Offsets in this data set are measured following the approach of Hodgkin et al. [85]. The maximum likelihood fitting procedures referred to throughout this section are summarised in appendix A.

A sub-sample of high signal-to-noise (S/N $\geq$ 10.0) point-like sources detected in both VIKING and the SDSS was defined and those sources with Vega like SDSS colours i.e. A0 stars with $u-g$, $g-r$, $r-i$ and $i-z$ in the range -0.1 to 0.1, were flagged. Figure 2.2

[1]Note that the VISTA photometric system is tied to, but distinct from 2MASS. See http://casu.ast.cam.ac.uk/surveys-projects/vista/technical/photometric-properties for further information on the calibration process.

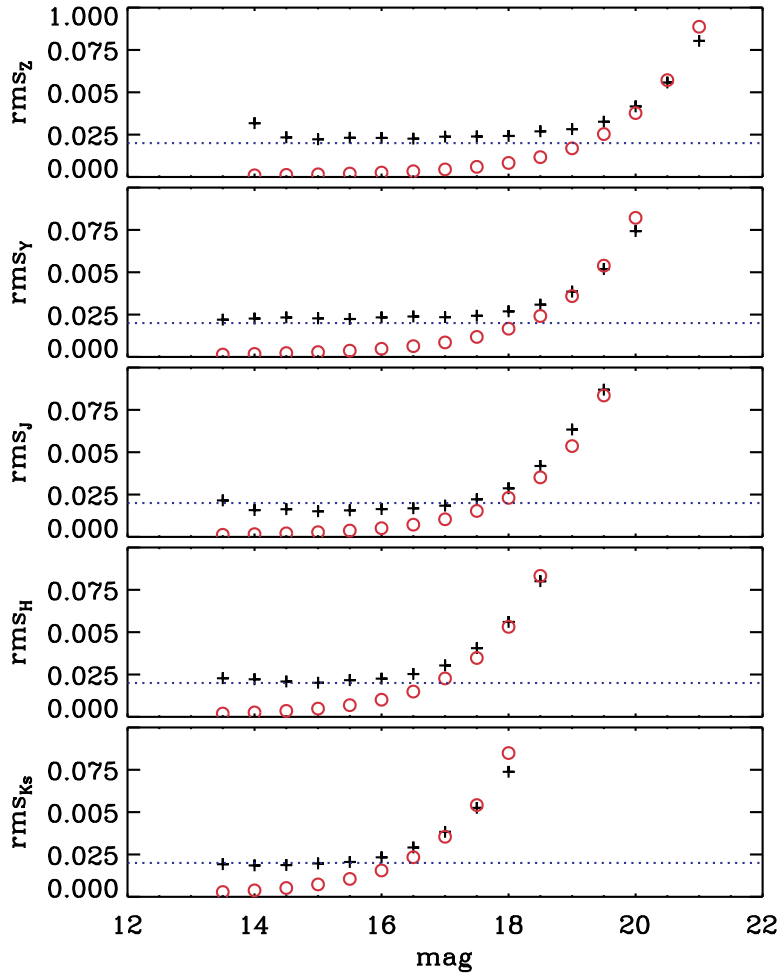


Figure 2.1: The rms residuals of repeat observations of VIKING sources in regions between dither offsets and tile overlaps. Crosses are rms residuals, while circles are pipeline errors which account for photon and sky background noise only. Pipeline errors underestimate residuals indicating some level of calibration error, which dominates at bright magnitudes. The dotted line indicates the survey goal, of 0.02 mag residuals for bright stars on photometric nights. The figure indicates that this goal is broadly achieved in the first full depth VIKING release.

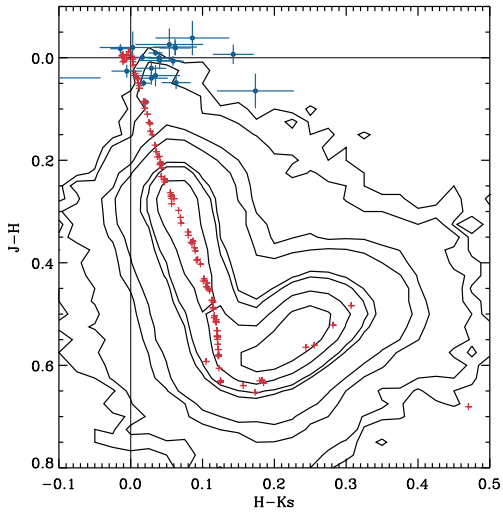


Figure 2.2: The VIKING J, H, Ks colour plane for high confidence stellar sources matched in VIKING and SDSS. Blue points are stars with Vega like colours in all SDSS passband combinations. Red crosses are synthetic photometry from the BPGS spectroscopic atlas [2].

plots the J, H and Ks colour-colour diagram for this set of sources. Blue points show the photometrically selected sample of A0 stars. Few of these objects are actually observed by VIKING since they are short lived and mostly dead in the Galactic halo. Where they are detected by VIKING in the Galactic plane they are invariably saturated. The small number of these objects highlights the difficulty of investigating the VIKING colours at this early stage. The median colours of this sample are $J - H = 0.005 \pm 0.038$ and $H - Ks = 0.033 \pm 0.023$ (where the median absolute deviation has been used to estimate the standard deviation), which perhaps hints at small offsets in one or both of H/Ks. Given that each of these colours has been directly calibrated from 2MASS standards, the analysis is continued assuming that any offset in these passbands is small. It will be shown shortly, that there is evidence that this assumption is a good one. Also plotted in figure 2.2 are red crosses, which show the synthetic colours of the stellar sequence derived from the BPGS spectroscopic atlas [2] (see chapter 3). The synthetic locus is in overall good agreement with the locus of bright stars in this colour-colour space, apart from a slight deviation at red $J - H$ colours. The source of this deviation is currently under investigation.

Following Hodgkin et al., the baselines on each Z, Y, J, H and Ks colour-colour combination are extended in one direction by plotting all possibilities against $u - Ks$. Offsets from zero are then measured from straight line fits to the data. To clarify this procedure the fits to each colour combination are shown in figure 2.3. Due to limited statistics our approach to the fitting differs slightly from Hodgkin et al. [85] who bin their data and take measurement errors based on standard deviations on binned median values. Conversely our approach is to fit to the photometrically selected A0 stars and retrieve errors

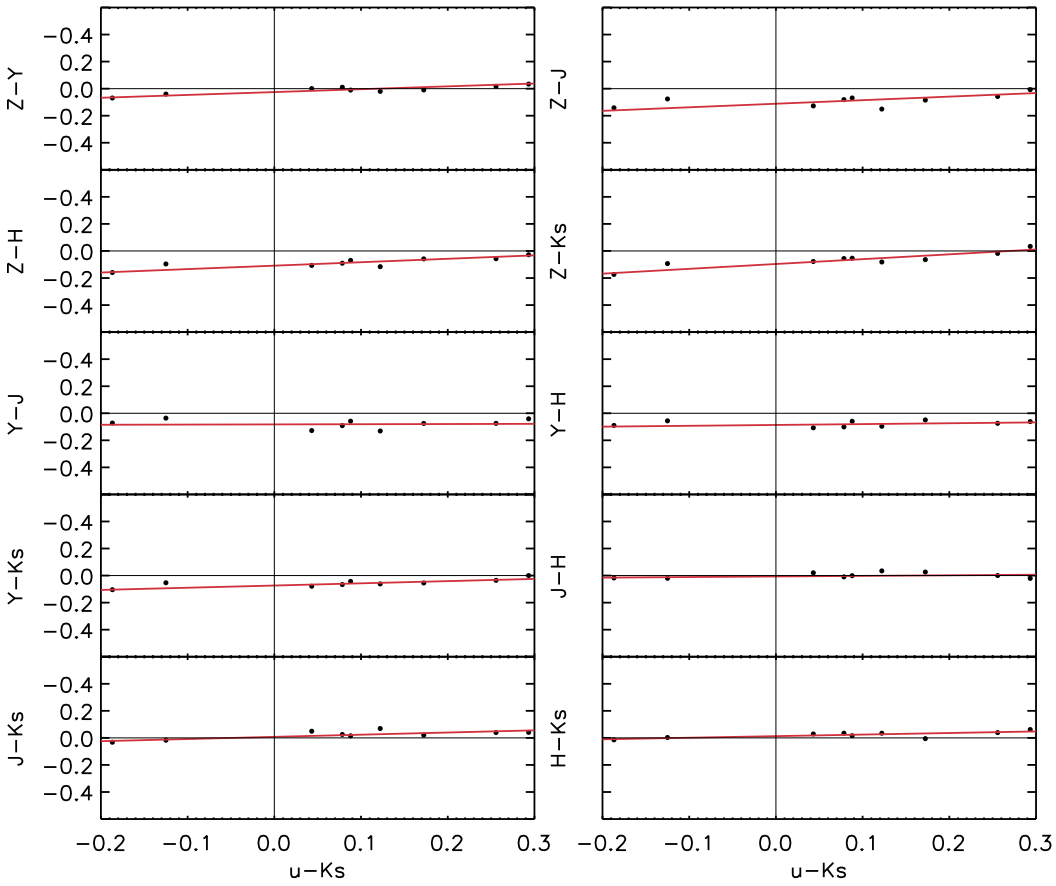


Figure 2.3: Maximum likelihood fits to stars with Vega like colours in SDSS with counterparts in VIKING. The results are summarised in table 2.1 and in the main text.

directly from the maximum likelihood fitting procedure. The results show appreciable offsets in the Z and Y bands only. These are summarised in table 2.1. The corresponding offsets found in the J, H and Ks bands are consistent with zero; $\Delta_{\rm JH} = 0.006 \pm 0.004$, $\Delta_{\rm JKs} = 0.008 \pm 0.005$, $\Delta_{\rm HKs} = 0.013 \pm 0.006$.

A further two results are shown in table 2.1, these were obtained by fitting to the blue end of the stellar locus from all the bright stars in the original selection. Again a straight line maximum likelihood fit is applied in the Z, J, Ks and Y, J, Ks colour-colour spaces and the offsets from the origins are measured. The best fitting lines are shown in figure 2.4 and the measured offsets in both the Z, J, Ks and Y, J, Ks colour spaces are summarised in table 2.1. In these cases the measured offsets are slightly larger than those measured in the VISTA-SDSS colour-colour spaces. These may well be attributed to small offsets in the J or Ks passbands as tentatively suggested by figure 2.2. Individual offsets to the J, H and Ks passbands implied by these fits are clearly smaller than those attributed to Z and Y and will not present significant problems in the following sections.

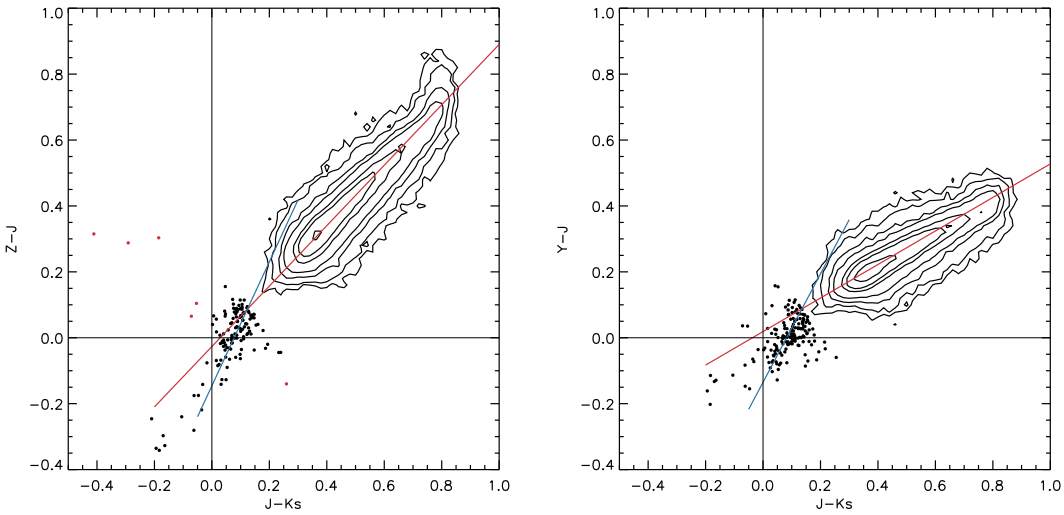


Figure 2.4: The VIKING Z, J, Ks and Y, J, Ks colour planes for high confidence stellar sources matched in VIKING and SDSS. A sample of blue stars is plotted with filled circles, obvious outliers were clipped from the sample and are shown in red. The blue line shows the maximum likelihood straight line fit to this distribution, while the red line shows the fit to the main sample. The Z and Y band offsets are measured from the blue star intercept to the origin.

	Δ_Z	Δ_Y
Δ_{ZJ}, Δ_{YJ}	-0.111 ± 0.002	-0.082 ± 0.002
Δ_{ZH}, Δ_{YH}	-0.109 ± 0.004	-0.086 ± 0.004
Δ_{ZKs}, Δ_{YKs}	-0.096 ± 0.006	-0.073 ± 0.005
Δ_{ZJKs}, Δ_{YJKs}	-0.146 ± 0.003	-0.135 ± 0.003
Mean	-0.116 ± 0.008	-0.094 ± 0.007

Table 2.1: Summary of Z and Y offsets from the Vega system.

VIKING-VIDEO

VIKING shares a region of partial overlap with two narrow field deep surveys; VIDEO and the Canada France Hawaii Telescope Survey (CFHTLS). VIDEO covers 12 $\deg^2$ over three fields in the sgp. The region of overlap with VIKING is $\sim$ 1 $\deg^2$ centred on the VIDEO tile at Ra $=$ 02h 26m 00s and dec $=$ -04° 30m 00s (the VIDEO-XMM3 field). Like VIKING, VIDEO observes in the ZYJHKs passbands, with 5σ limiting depths of Z=25.2, Y=24.0, J=23.6, H=22.6 and Ks=21.7, corresponding to an approximate $\sim$ 2.5 mag gain over VIKING.

CFHTLS was carried out between 2003 - 09 with joint Canadian-French grey and dark time on the wide field imaging camera, MegaCam. The survey includes a number of fields varying in depth and area. In particular four deep fields were imaged in SDSS-like passbands to 5σ limiting depths of $u_{\rm CFH} = 26.6$, $g_{\rm CFH} = 28.0$, $r_{\rm CFH} = 27.6$, $i_{\rm CFH} = 27.0$ and $z_{\rm CFH} = 25.7$. The Deep field 1, shares $\sim$ 1 $\deg^2$ of overlap with VIKING and VIDEO.

The deep optical and infrared imaging of VIKING sources in this region, is a valuable commodity that may be put to good use in investigating the nature of very faint VIKING sources. As we shall see in chapter 3, combinations of optical and infrared colours allow very accurate star-galaxy classifications to be made. VIKING sources were matched to VIDEO sources and then to CFHTLS sources producing a nine-band optical to near-infrared data set. The resulting catalogue is hereafter referred to as the VVC catalogue. The exact relevance of this data set will become clearer in subsequent sections.

The VIDEO catalogue was produced by a team at the University of Hertfordshire whereas the VIKING catalogue has been produced by the VDFS. Differing reduction procedures have likely led to small systematic offsets between the two data sets. In order to make comparisons between the two catalogues any systematics need to be accounted for i.e. the two catalogues need to be placed on a the same photometric system. At the time of writing the Z-band VIDEO photometry is unavailable, fortunately Y and J photometry is sufficient for applications in this thesis. Figure 2.5 plots offsets in VIKING-VIDEO YJ photometry against VIKING Y and J for stars in the VVC catalogue (see chapter 3). Offsets from zero are clear and appear to be close to linear at least for bright stars of Y or J $\leq$ 19. Straight line fits to each of the plots between 15 $\leq$ Y $\leq$ 19 and 15 $\leq$ J $\leq$ 19 show slopes consistent with zero. The fitting is summarised as follows $\rm Y_{VIKING} - Y_{VIDEO} = -0.38$ and $\rm Y_{VIKING} - Y_{VIDEO} = -0.29$.

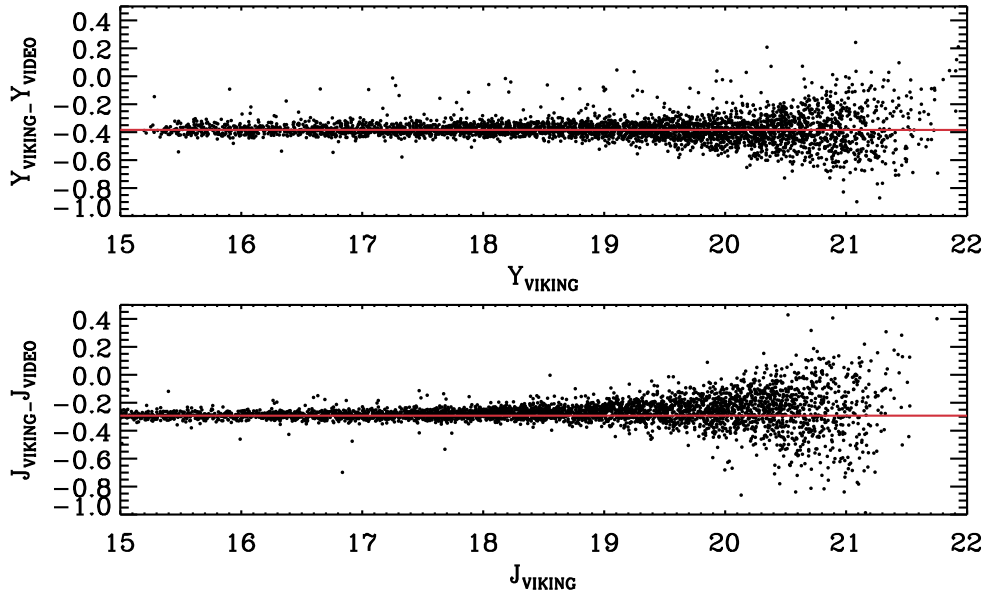


Figure 2.5: Offsets between VIKING and VIDEO photometry for stars in the VVC catalogue due to differences in reduction procedures including default aperture radii. Straight line fits to the data indicate zero slope and offsets of the form $Y_{VIKING} - Y_{VIDEO} = -0.38$ and $Y_{VIKING} - Y_{VIDEO} = -0.29$. A description of the fitting procedure is given by Williams et al. [3]

2.4 Bright star halos

Two data quality issues arise near the vicinity of bright stars. The first is that a star will produce diffraction spikes caused by a 'tree-ring' like pattern of small imperfections on the primary mirror. By some complicated optical process, local maxima are produced and are interpreted as detections by the CASU source extractor. In most cases the detections have large ellipticities and are flagged as extended as in figure 2.6, which shows detections around the same bright source in three different bands. Extended sources are plotted in green and point-like sources are shown in blue.

Since these events are not real detections they are not matched across bands and are easily removed where the particular science application requires multi-band photometry. Inevitably though a fraction of these artifacts are falsely matched between bands and can produce spurious sources in an otherwise well constrained catalogue.

The second issue is regarding real sources with spurious photometry. When a source is in the near vicinity of a bright star, its photometry may be affected by an overlapping or near coincident diffraction spike or more commonly by a ghost image of the star produced by the various reflections within the filter and between the filter and the detector. Examples of such incidents are shown in figure 2.7 (a).

The length of diffraction spikes and the extent of ghosting are strong functions of brightness of the star. However in the case of the former the complex nature of the

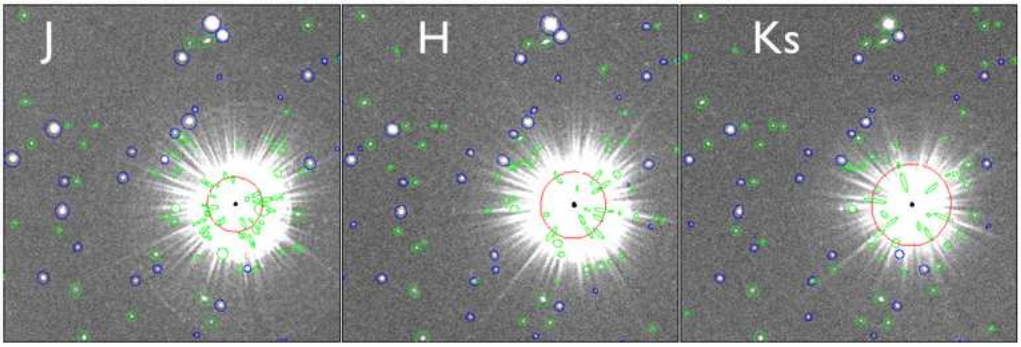


Figure 2.6: Example of false detections of local maxima produced around the same bright star in the J, H and Ks bands. Detections are indicated by blue for point-like objects, green for extended objects and red for saturated objects. Most false detections in the vicinity of the bright star are flagged as extended and may be removed by querying the VSA for single band detections of extended sources in the vicinity of bright stars. A fraction of spatially coincident maxima are falsely matched between bands and these are more difficult to remove, potentially contaminating otherwise well constrained catalogues (credit: CASU).

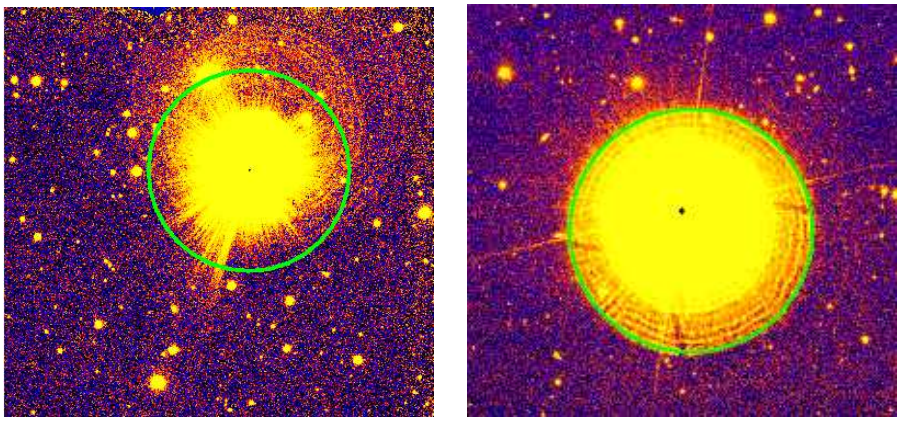
diffraction pattern means that diffraction spike length is not a simple function of brightness but depends further of the position of the star relative the optical axis of the telescope. Consequently the most straight forward way to remove objects affected by these appendages is to define a region around a bright star which encloses the vast majority of the area affected by the diffraction as shown by the highlighted region in figure 2.7 (a). Objects that fall into this region can then be flagged for visual inspection.

To this end, figure 2.8 shows the radii of circular regions affected by diffraction as a function of J-band magnitude. The use of 2MASS over VISTA magnitudes is necessary to accurately determine the magnitudes of bright stars since the VIRCAM chips saturate at $J \lesssim 13$, whereas the response of the 2MASS instrument remains linear at much brighter magnitudes. This additional requirement means that VISTA catalogues must be matched to the 2MASS point source catalogue (PSC). The VSA facilitates this via some simple SQL since it stores the 2MASS PSC locally.

Regions around bright stars effected by diffraction spikes were measured visually in the astronomical imaging and data visualisation application SAOImage DS9[2], by adjusting the radius of a circular region placed on the sky coordinates of the stellar flux peak as inferred from 2MASS measurements. Given the obviously subjective nature of this task, there is a surprisingly well defined trend of increasing radius with brightness visible within the measurement scatter. The trend is fit to the second order polynomial with the IDL[3] routine SVDFIT; $r = 0.0004\,(J_{2MASS})^2 - 0.0096\,(J_{2MASS}) + 0.0671$, as shown by the red curve.

With an empirical function describing the extent of the of diffraction spikes from the

[2]http://hea-www.harvard.edu/RD/ds9/
[3]http://www.exelisvis.com/ProductsServices/IDL.aspx



(a) (b)

Figure 2.7: Panel (a) shows extensive diffraction related artifacts around a bright star. These artifacts can produce false maxima or perturb the photometry of real sources. Consequently detections close to bright stars need to be visually inspected to clarify the reliability of their photometry. The green region shows an area in which detections are flagged. A number of these regions were measured in order determine an approximate empirical relationship between the extent of diffraction spikes and stellar brightness (see text for details). Panel (b) shows ghosting around a bright star caused by various reflections between the filter and detector interfaces. The nature of these ghosts are less complicated than diffraction and their extent and position relative to the associated bright star can be easily modelled. The green region shows the extent of the largest ghost seen in VIKING images.

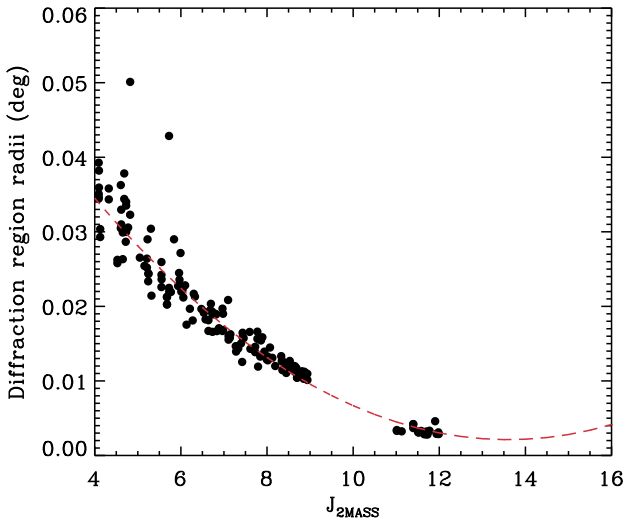


Figure 2.8: Radii of regions around bright stars enclosing the majority of diffraction related artifacts. A few outliers are present from the main distribution but in general the scatter is small given the subjective nature of the measurements. The distribution is fit to a second order polynomial; $r = 0.0004\,(\mathrm{J_{2MASS}})^2 - 0.0096\,(\mathrm{J_{2MASS}}) + 0.0671$. By J = 12 diffraction ceases to be a problem around most bright stars.

intensity peaks of bright stars it is possible to automate the flagging of sources with potentially spurious photometry via their proximity to the intensity peaks of bright 2MASS stars in VIKING.

A similar approach can be pursued to flag objects in the ghosts of bright stars caused by reflections between various interfaces in the filter and detector architecture. In general the ghost pattern is formed of three different reflection combinations forming strongly defocused images of varying intensity and each are offset from the real image of the star. This is illustrated in figure 2.7 (b). For the brightest stars, such as that in the figure, all three ghosts may be visible above the background if not obscured by the starlight. In the case of figure 2.7 (b) at least two ghosts are visible. As stars become fainter the ghosts begin to disappear. In order to understand the ghost pattern and its dependency on brightness it is necessary to look at the optical design of the VISTA telescope.

Figure 2.9 is a ray diagram superimposed on a simplified schematic of the optical system. After reflections from the primary and secondary mirrors, M1 and M2, light is directed through a vacuum window, WINDOW, after which lie three corrector lenses, L1, L2 and L3, which direct the light through the filter to the focal point on the detectors. The diagram shows the generation of three different images over the field of view. In 3-dimensional reality, each image would be focused onto the detector at the apex of a conical volume. In this 2-dimensional representation three rays are traced through the optical system. The central ray, known as the chief ray, never meets the detector since it is obstructed by the secondary mirror. The corrector lenses imply that the angle between each of the rays originating from the extrema of the field and the chief ray is 8.75 deg and further that the angle between the chief ray and the optical axis, i.e. the non-telecentricity, is $\phi_1^c = 5.9$ deg for a star at the corner of the field i.e. 0.825 deg off-axis (half the field diameter). Assuming linearity ϕ_1^c is then in general a factor $5.9/0.825 = 7.15$ the star off-axis angle. In the remainder of this chapter, the characteristics of the optical system are used to develop a model to describe the physical position and extent of the Ghost pattern on the detector array.

Figure 2.10 zooms in on the focus of the blue rays *a*, *b* and *c* in figure 2.9, tracing out the origin of the smallest ghost via two reflections within the filter. The filter of thickness h_2, is suspended in a vacuum a distance h_3 above the detector plane, which is without 15 mm for all passbands. The reflected ray paths are depicted by solid black and blue lines for rays *a* and *b* respectively. The paths taken by unreflected rays to the focus are shown by the thick grey lines and the chief ray is shown by the black dashed line.

The distance d_{small} is the diameter of the small ghost, which is equivalent to the extra path length travelled by each ray in the direction parallel to the detector plane as a result of the reflections.

The angles ϕ_1^a, ϕ_2^a are related by Snell's law and the small angle approximation as $\phi_1^a \simeq$

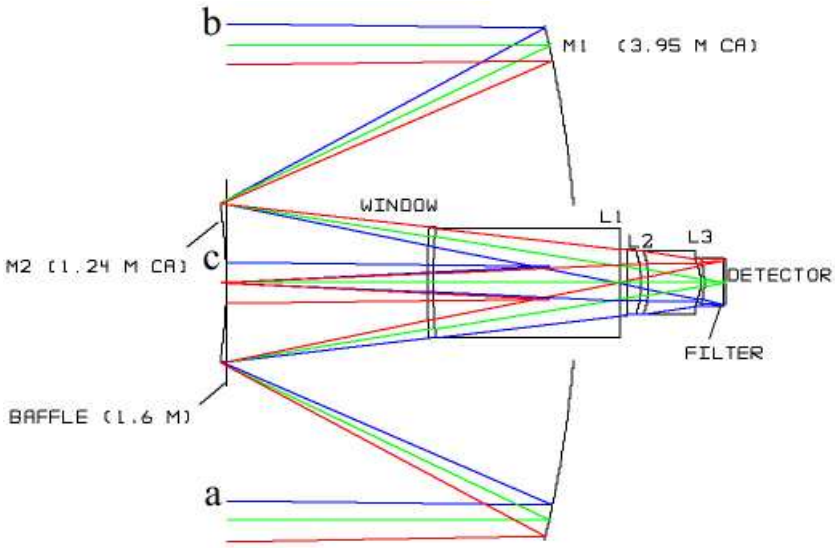


Figure 2.9: The path of three light rays, *a*, *b* and the hypothetical chief ray *c*, through the optics of the telescope to focus on the detector are shown in blue. Similar paths are shown in red and green which focus images elsewhere is the field. The point at which the blue ray passes through the filter is blown up in figures 2.10 and 2.11 (credit: VISTA; adapted from original.)

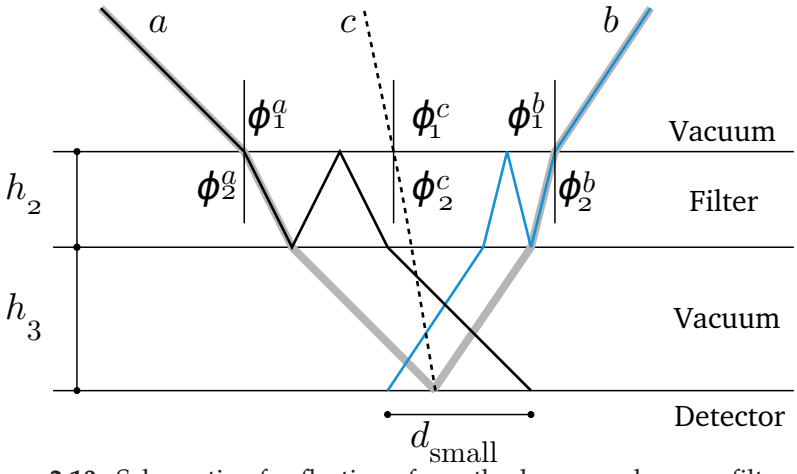


Figure 2.10: Schematic of reflections from the lower and upper filter-vacuum interfaces, which produce the smallest of three circular ghost patterns around bright stars. The unreflected paths of rays *a* and *b* are shown by thick grey lines. A fraction of light is reflected from the detector and then off the filter. The reflected rays are shown as black and blue solid lines for *a* and *b* respectively. The dashed line shows the hypothetical path of the chief ray *c* which in reality does not make it to the detector since its path is blocked be the secondary mirror. *a* and *b* are offset from *c* by an angle of 8.75 deg. The various angles denoted ϕ show the offsets between each of the rays and the pointing direction of the telescope, which is at all times normal to the filter and detectors. d_{small} is the diameter of the ghost.

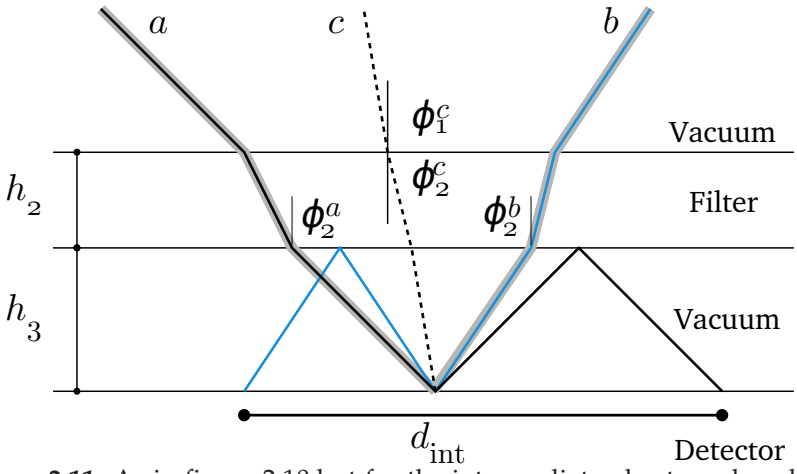


Figure 2.11: As in figure 2.10 but for the intermediate ghost produced by reflections from the detector and lower filter-vacuum interface.

$\eta\,\phi_2^a$ and similarly $\phi_1^b \simeq \eta\,\phi_2^b$, where η is the refractive index of the the filter. Some straight forward geometry gives

$$d_{\text{small}} \simeq 2\,h_2\left[\tan\left(\frac{\phi_1^a}{\eta}\right) + \tan\left(\frac{\phi_1^b}{\eta}\right)\right] \tag{2.4.1}$$

Since rays a and b are offset from the chief ray c by 8.75 deg, then $\phi_1^a = \phi_1^c + 8.75$ and $\phi_1^b = 8.75 - \phi_1^c$. Converting to radians, substituting into equation 2.4.1 and invoking the small angle approximation gives,

$$d_{\text{small}} \simeq \frac{h_2\,0.611}{\eta}. \tag{2.4.2}$$

The extent of the ghost is thus constant over the field.

The intermediate ghost is produced by a double reflection off the detector and the lower vacuum-filter interface as shown in figure 2.11. A similar argument as given in the case of the small ghost yields,

$$d_{\text{int}} \simeq h_3\,0.611 \tag{2.4.3}$$

The largest ghost is a result of the combination of reflections that produced the small and intermediate ghosts i.e. a reflection off the detector and a reflection off the upper filter-vacuum interface. Thus the extent of the resultant ghost is a combination of those of the two smaller ghosts,

$$d_{\text{large}} \simeq d_{\text{small}} + d_{\text{int}} = 0.611\left(\frac{h_2}{\eta} + h_3\right). \tag{2.4.4}$$

All this is centred on the hypothetical path of the chief ray, so by tracing the chief ray

Filter	Thickness (mm)	Refractive index (100 K)
Y	9.656	1.508
J	9.770	1.505
H	10.00	1.499
Ks	10.37	1.436

Table 2.2: Attributes of each filter necessary for calculating the extent and offset of each ghost (Sutherland; private communication). At the time of writing Z filter data is unavailable. All filters are offset a distance 15 mm above the detector.

through each of the above reflections, the difference between the reflected and unreflected paths of the chief ray in the direction parallel to the detector gives the offset of each ghost from the star. The chief ray is subject to the same reflections as a and b, and the offsets from the centre of the focused image are

$$\Delta d_{\rm small} \simeq \frac{2\,h_2\,\phi_1^c}{\eta} \tag{2.4.5}$$

$$\Delta d_{\rm int} \simeq 2\,h_3\,\phi_1^c \tag{2.4.6}$$

$$\Delta d_{\rm large} \simeq 2\,\phi_1^c\left(\frac{h_2}{\eta} + h_3\right). \tag{2.4.7}$$

Unlike the ghost diameters, the extent of the offsets are a function of the star off-axis angle.

The filter dimensions and their refractive indices are given in table 2.2. In order to convert the output of the above equations into an on sky distance, the VISTA plate-scale is required. Ignoring field distortion[4] the relevant value is 8.281×10^{-5} radians/mm.

The next step is to determine the visibility of the ghosting above the background as a function of star brightness. Like the diffraction spike analysis, the most straight forward way to do this is visual inspection of a number of image cutouts centred on bright stars. Figure 2.12 plots 2MASS J-band magnitude versus largest visible ghost size. Again scatter about the theoretical ghost radii is minimal although there are a number of outliers, which highlight the subjective nature of the task.

In flagging objects affected by diffraction spikes or ghosting, it is important to remain conservative, defining a slightly larger zone around a bright star than is perhaps necessary. All flagged objects can then undergo visual inspection and those clearly affected by these issues can be discarded.

[4]http://casu.ast.cam.ac.uk/surveys-projects/vista/technical/astrometric-properties

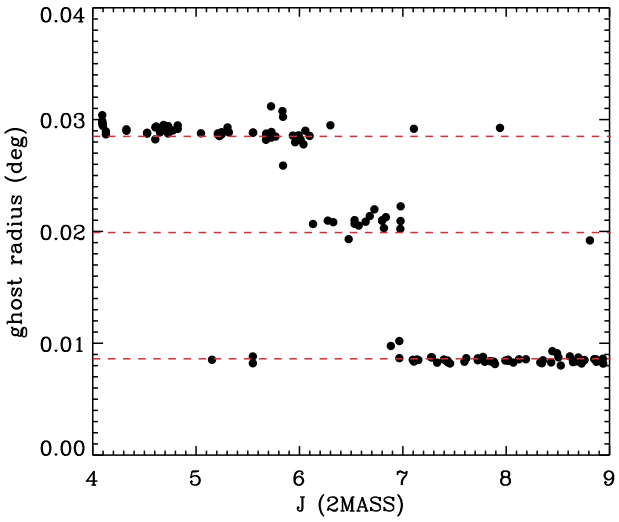


Figure 2.12: The extent of ghosts around bright stars. Points are measured direct from VIKING imaging frames, red dashed lines show theoretical predictions. The largest ghost appears until $J_{2MASS} \lesssim 6$ the intermediate $J_{2MASS} \lesssim 7$ and the small $J_{2MASS} \lesssim 9$. Theoretical predictions agree well with measured values but outliers are present due to the subjective nature of the measurements. The best approach in flagging sources around bright stars is to remain conservative, defining larger regions than perhaps is necessary.

2.5 Summary

In this chapter a number quality issues are investigated. In summary the main results are as follows;

- Re-measurements of stellar sources are extracted from the VSA by querying for objects in the overlap regions between frames and tiles. Root mean square residuals indicate that the VIKING survey broadly conforms to the nominal survey requirements which ask for a 2 per cent uncertainty in the measurement of bright stars in all passbands. Uncertainty is shown to be dominated by calibration noise at bright magnitudes.

- The offset of the VISTA photometric system from Vega is investigated. This is an issue because VISTA photometry is calibrated from 2MASS standards and linear extrapolations of VISTA-2MASS colour equations from JHKs wavelengths are used to calibrate the ZY photometry. In comparing colours of blue and red stars in various VISTA-SDSS colour spaces significant offsets from the Vega system are determined in the Z and Y bands.

- Data artifacts and spurious photometry are found to be prevalent in the regions around bright stars. These issues are brought on by diffraction and image ghosting. A set of recipes to determine the extent of affected regions as a function of

the bright star sky-coordinates are presented, allowing for the flagging and subsequent visual inspection of affected sources.

These issues are relevant to high-z quasar searches in VIKING; ensuring the quality of the catalogue photometry has obvious benefits for any science application. In following chapters, Vega zeroed photometry is synthesised in order to determine the properties of both quasars and stars in VISTA colour spaces, therefore it is important that offsets from Vega are known and well measured. Objects in the halos of bright stars cause spurious detections and perturb the photometry of real sources leading to unwanted contamination. The recipes presented here will be used in following chapters to remove contamination from quasar candidate lists.

The work presented here is also useful to the users of VIKING data in more general science applications.

CHAPTER 3

The cool star, galaxy and quasar populations

VISTA's transmission curves are shown in figure 3.1. These include the relevant contributions from the instrument and detector, mirror reflectivity and transmissions through the atmosphere and filters under typical photometric conditions. Also plotted is the Large Bright Quasar Survey's [LBQS; 63] template quasar spectrum [86], redshifted to $z = 6.8$ and with the characteristic HI absorption trough artificially imposed blueward of redshifted Lyα at $\lambda_{\mathrm{Ly}\alpha} = 7.8 \times 1216$ Å. The spectra of cool M8- and T5-dwarf stars are plotted alongside the quasar spectrum to highlight, the recognised difficulty of differentiating between high-z quasars and cool stars at the foot of the main sequence and below the hydrogen burning cut-off.

M-dwarfs out-number high-z quasars by at least a factor $\sim 10^5$ (for instance Bochanski et al. [87], obtained a volume limited sample of $\sim 15 \times 10^6$ SDSS objects with M-dwarf like colours at $z < 21.2$. Over a similar depth and area there are currently 19 SDSS quasars with $z \sim 6$ reported by Fan et al. [11]), while the cooler L and T spectral types are by all accounts, a lot more rare [88, 89]. The broad similarities between high-z quasars and M-dwarf spectra, coupled with their relative surface densities means that most colour selected high-z quasar candidates are in fact M-dwarfs with scattered quasar like colours resulting from large photometric errors. The function of a good colour selection criterion then, is to maintain a large selection space while simultaneously minimising contamination from false positives and optimising the completeness of the search.

Precision selection strategies can be developed with a detailed understanding of the error perturbed magnitude and colour distributions of the target and contaminant populations, both as a function of redshift and direction on the sky. These strategies are developed with the accumulation of large data volumes. The first selection and photometric followup of VIKING candidates was scheduled for summer 2011, at which point in excess of $\sim 300\,\mathrm{deg}^2$ of science ready data was expected to be available. Time

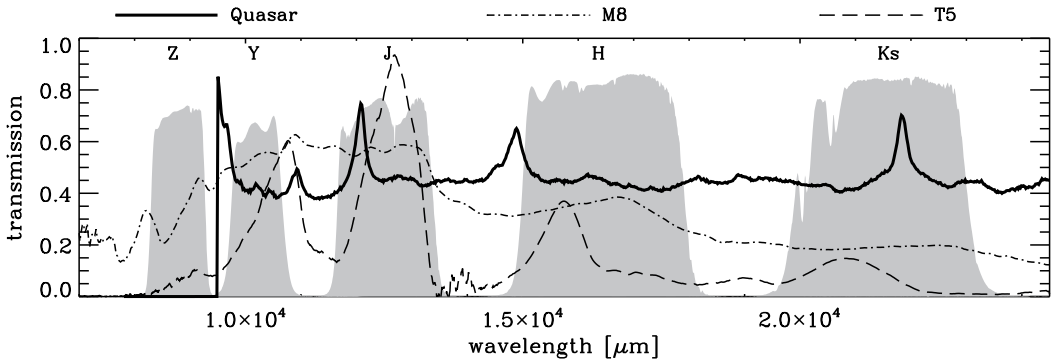


Figure 3.1: Transmission curves for the VISTA ZY JHKs photometric system including wavelength dependent losses are labelled and shaded grey. Also shown is the LBQS template quasar spectrum artificially redshifted to $z = 6.8$ and with the characteristic absorption trough imposed blueward of redshifted Lyα (solid black curve). For comparison the spectra of cool M8 and T5 dwarfs are plotted (dash-dot and dashed curves respectively).

constraints demanded that a selection strategy be in place well before this time and consequently initial candidate selection was based on limited available data in initial releases. Thus, the task of optimising selection constraints to provide a complete, efficient and non-biased search necessitated a comprehensive assessment of the initial data, combined with detailed modelling of the observations both as a function of redshift and direction on the sky [e.g. 90].

In the following chapter a model of the magnitude, colour and spatial distributions of the cool star and quasar populations is developed and assessed along with $\simeq 200\,\mathrm{deg}^2$ of initial imaging data to accurately determine the nature of target and contaminant populations in quasar searches over the VIKING field.

Henceforth, all quoted photometry is defined on the Vega system, such that Vega is a zero magnitude star in all (relevant) passbands. Unless specifically stated otherwise, magnitudes are quoted in the traditional logarithmic form except for the those in the VISTA Z band, which are quoted in the asinh form [91]. asinh magnitudes take the same values as the equivalent logarithmic form for high signal-to-noise measurements, but have the advantage of linear dependence on flux at low signal-to-noise or even negative flux, which will be the case for most $z \gtrsim 6.5$ quasars. A softening parameter, denoted b, is required to define the scale at which linear dependence dominates. For the Z band we choose b equivalent to $Z = 24.1$, which gives the asinh magnitude of an object with zero flux and is set approximately equal to the sky noise seen in the initial data frames.

3.1 Synthetic photometry and characteristic colours

To obtain representative colour and magnitude distributions of high-z quasars and cool stars it is necessary to first characterise the flux distributions of a distinct set of representative objects. The method of synthetic photometry provides a precise means to computationally measure the flux distribution of any object with a sufficiently well sampled spectrum. For a given source the basic premise is the definition of the mean flux density at the top of the Earth's atmosphere

$$f_{\rm i}(\Gamma_{\rm i}) = \frac{\int \Gamma_{\rm i}(\lambda) f_\lambda(\lambda) \lambda \, {\rm d}\lambda}{\int \Gamma_{\rm i}(\lambda) \lambda \, {\rm d}\lambda}, \tag{3.1.1}$$

where $\Gamma_{\rm i}(\lambda)$ is the dimensionless response function of the photometric passband i. The function $\Gamma_{\rm i}(\lambda)$ is in fact the product of a number of wavelength dependent quantities which describe the signal attenuation resulting from the filter transmission, atmospheric extinction, mirror reflectivity and instrument and detector response. The factor λ converts the energy flux to a photon flux as appropriate for modern photon counting detectors, where the flux density integrated over the bandpass i is proportional to the observed photon count rate.

With a knowledge of both $\Gamma_{\rm i}(\lambda)$ and the object's spectrum, $f_\lambda(\lambda)$, at satisfactory resolutions, equation 3.1.1 can be evaluated numerically. Thus the process is differential in nature and mirrors the procedure undertaken during real photometric observations. Corresponding magnitudes can then be computed via the usual definitions

$$m_{\rm i} = -2.5 \log_{10} f_{\rm i}(\Gamma_{\rm i}) + p_{\rm i}, \tag{3.1.2}$$

$$m'_{\rm i} = -1.08574 \left[{\rm asinh} \frac{f_{\rm i}(\Gamma_{\rm i})}{2{\rm b}} + \ln {\rm b} \right] + p'_{\rm i}, \tag{3.1.3}$$

where $m_{\rm i}$ and $m'_{\rm i}$ are the logarithmic and asinh magnitudes defined with reference to the bandpass specific constants $p_{\rm i}$ and $p'_{\rm i}$, i.e. the zero-point magnitudes defined by the relevant photometric system in the arbitrary bandpass i.

In this section, characteristic colours are computed by evaluating equations 3.1.1-3.1.3 for a number of quasar and cool star spectra in the VISTA passbands. The photometric zero-point is defined by the spectrophotometric standard star Vega (α-Lyrae) and is calculated by evaluating equation 3.1.1 for a suitable Vega spectrum. Indeed the choice of spectrum is an important one, as we saw in Chapter 2 it is not uncommon to find small offsets between synthetic and observed photometry, which are generally resolved with the acquisition of large data sets from which accurate zero points can be inferred.

Vega is a near-pole-on fast rotator [92], whose surface temperature ranges from 7900K

at the equator to 10150K at the poles. Accurate measurement of the absolute flux distribution of Vega is therefore notoriously difficult [93, 94] particularly between the optical and near-infrared. The application of an accurate reference SED is thus an important step in the processes described above. Widely used Vega SEDs include those of Cohen et al. [95], and Bohlin and Gilliland [96], the latter of which has recently undergone revision [97]. The agreement between the Cohen et al. [95] and Bohlin [97] SEDs is inside the 1 per cent level in the near-infrared, the latter being precise to within 1-2 per cent. All synthetic photometry found in this thesis use the Vega reference SED of Bohlin [97], such that Vega is a zero magnitude star in all VISTA passbands.

3.1.1 High-redshift quasars

With the above algorithm in hand the question is how to obtain a sample of high-z quasar spectra on which to apply it. With VIKING, we are targeting quasars at $z \geq 6.5$, a domain in which (at the time of writing) there is just one known example [37]. Of the fifty or so known at $z \simeq 6$ only a few have been spectroscopically observed with at least moderate resolution and coverage in the VISTA ZYJ bands. So most examples of the known quasars at high-z are not useful. Rather, it is better to consider objects at lower redshift where space densities are higher and sampling is more complete.

Recent comparisons between high and low-z quasar spectra have shown no obvious evidence for metallicity evolution in the rest-frame ultra-violet SEDs [37, 49, 50, although see [55] for possible evolution of dust properties]. In essence, the rest frame ultra-violet properties of high- and low-z quasars are generally indistinguishable and it is a reasonable approximation to simulate intrinsic quasar SEDs at $z \geq 6.5$ by artificially redshifting a low-z SDSS sample.

Significant differences between $z \gtrsim 6.5$ quasars and those at lower redshift are apparent shortward of Lyα due to increasingly strong absorption in the intergalactic medium with redshift (see also chapter 1). After artificially redshifting a low-z sample it is necessary to correct for this effect. At $z = 6.5$ the optical depth of the intergalactic medium to Lyα photons is $\gg 1$ [e.g. 11]. For the same neutral density the optical depths to Lyβ and Lyγ are factors of 6.2 and 17.9 less than for that of Lyα, but their line strengths are comparatively small and they can be reasonably neglected [e.g. 98]. Transmission shortward of $\lambda = 1216\ (1 + z)$ Å is thus effectively nil, and absorption by intergalactic hydrogen can be adequately modelled in $z \geq 6.5$ quasar spectra by setting transmission shortward of redshifted Lyα to zero.

Low-z optical quasar spectra that sample the rest-frame wavelengths of $6.5 \lesssim z \lesssim 7.5$ quasars in VISTA's ZYJ filters are available from the SDSS Quasar Catalog. IV. [99, recently superseded]. This repository contains the spectra of tens of thousands of quasars to redshifts of ~ 5. With the SDSS spectrographs, the required redshift space sampling can be achieved by redshifting quasars in the range $3.1 \lesssim z \lesssim 3.2$. Over this interval

the SDSS selection algorithm suffers from minimal stellar contamination and is highly complete [100].

3.1.2 Cool stars

Cool Galactic stars, specifically mid-M- to late T-dwarfs, are well known to be the most prominent Galactic contaminants of optical and near-infrared selected samples of high-z quasar candidates [101, 102]. Stars of spectral type O to K pose less of a problem, since they are hot enough that their colour sequences are highly temperature regulated. The coolest brown dwarfs are by all accounts rare [88, 103], but they are at least more common than $z \geq 6.5$ quasars and many examples of moderate to high resolution near-infrared spectra, right up to the coolest known dwarfs, are publicly available.

A library of cool star spectra has been collected from the literature covering VISTA's Z, Y, J, H and Ks passbands, these include all 21 M-, 30 L- and 22 T-dwarf spectra on which Hewett et al. [72] conducted synthetic photometry in the WFCAM passbands. References for these objects can be found in the above paper.

The library is further supplemented by spectra obtained from the DwarfArchives[1], the IRTF spectral library[2], the SpeX Prism spectral library[3], the Keck LRIS spectral library[4] and Sandy Leggett's L- and T-dwarf archives[5]. At this stage it is useful to review the general approach to near-infrared high-z quasar selection in visual form. In figure 3.2 a ZYJ colour-colour diagram is presented with the derived synthetic photometry. The quasar tracks are colour coded according to their redshift in steps of $\Delta z = 0.01$ and the average track is labelled at various redshifts for clarity. The box bounded by the dashed line illustrates the general principal; the box includes almost all $6.5 < z < 7.5$ quasars, rejecting typical E1 galaxies [104] and almost all foreground stars. The position of each quasar in $Z - Y$ is a function of redshift, reflecting the movement of the Lyα break through the Z-band. The $Y - J$ colours have only minor dependence on redshift until $z = 6.9$ where the locus begins to turn over as Lyα enters the observed Y band.

Conversely, the distribution of stellar colours is driven largely by temperature, which acts to redden spectral types approaching early L. In later types, behaviour is more complicated as clouds act to mask spectral features and redden near-IR colours [see 105]. Pressure broadened NaI and KI absorption features begin to appear at mid to late L, etching away at flux in the Z band. The $Z - Y$ colours are forced rapidly redward into the T sequence characterised by the appearance of CH_4.

The simple cuts outlined above will need to be modified for a real survey, since ran-

[1] http://spider.ipac.caltech.edu/staff/davy/ARCHIVE/links.shtml
[2] http://irtfweb.ifa.hawaii.edu/~spex/IRTF_Spectral_Library/
[3] http://pono.ucsd.edu/~adam/browndwarfs/spexprism/
[4] http://www.stsci.edu/~inr/ultracool.html
[5] http://staff.gemini.edu/~sleggett/LTdata.html

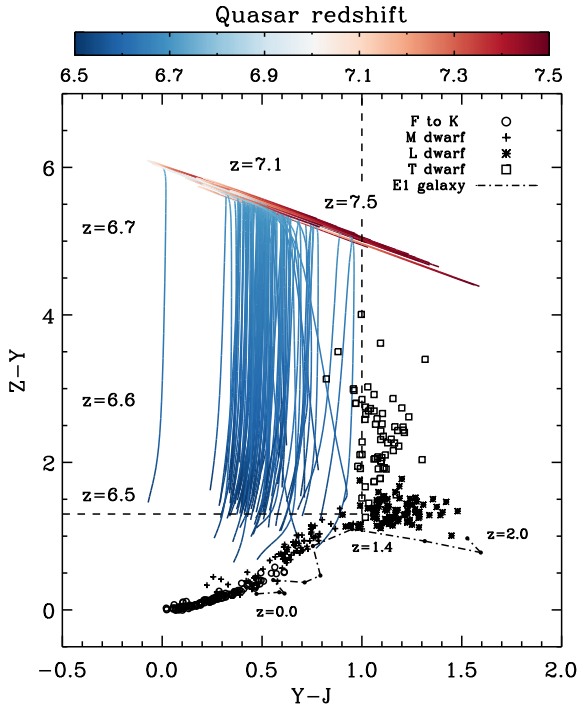


Figure 3.2: The Z, Y, J colour plane; high-z quasar redshift evolution tracks are colour coded as indicated by the colour bar and extend from z = 6.5-7.5. Star and galaxy colours are indicated by the legend. The dashed line describes the general principals of quasar colour selection in the Z, Y, J passbands, enclosing the vast majority of high-z quasars while rejecting the vast majority of galaxies and foreground stars.

dom photometric errors will broaden the measured loci, and thus can scatter stars and galaxies into the quasar selection region; since all stars and galaxies outnumber genuine high-z quasars by a factor $> 10^6$ ($J \lesssim 21$), a realistic survey will need to ensure that the cuts are sufficiently far from the stellar locus to reject almost all stars and galaxies.

3.2 Simulating the VIKING catalogue

Having used synthetic photometry to predict the colours of cool stars and quasars in the VISTA passbands, the next step is to combine these with models of cool-star and quasar number counts, with the addition of photometric scatter, to produce a simulation of VIKING multicolour photometry over the entire survey.

3.2.1 Quasar number counts

Since the majority of followup candidates from VIKING will consist mainly of M-dwarfs, other cool stars and galaxies the specific number of quasars expected from VIKING is of less relevance than their photometric properties. Nevertheless, the calculation is important for two other reasons; firstly to demonstrate that followup at this stage presents a statistically viable prospect of discovering a quasar and secondly because doing so will allow qualitative constraints to be placed on the luminosity function when the VIKING quasar search begins to bare fruit.

The rationale behind much of the work undertaken by previous surveys has been to describe the evolution of the quasar population as a function of look-back time. These surveys have provided measurements of the surface density of quasars as a function of their brightness and redshift. Converting redshift to distance, the surface density can be deprojected to give the comoving density of quasars within some luminosity interval. This fundamental quantity is the quasar luminosity function (QLF) discussed in chapter 1.

The latest determination of the QLF at $z \simeq 6$ was undertaken by Willott et al. [4] and combines discoveries from the CFHQS with the more luminous SDSS main and deep samples. The input catalogue comprises of 40 quasars sampling the redshift range $5.74 < z < 6.42$. The bright end of the binned luminosity function is well constrained to a power law with some evidence for a flattening in the slope provided by a single quasar at lower luminosities. In keeping with work at lower redshifts, where the QLF is well constrained over a large magnitude interval [106, 107], the Willott et al. parametric function follows a double power law of the form,

$$\Phi(M_{1450}, z) = \frac{10^{k(z-6)}\Phi(M^*_{1450})}{10^{0.4(\alpha+1)(M_{1450}-M^*_{1450})} + 10^{0.4(\beta+1)(M_{1450}-M^*_{1450})}}, \tag{3.2.1}$$

Table 3.1: Parameters of the z = 6 luminosity function of Willott et al. [4]

Density evolution parameter	k	-0.47
Normalisation	$\Phi(M_{1450}, z)$	$1.14 \times 10^{-8}\,\mathrm{Mpc}^{-3}\,\mathrm{mag}^{-1}$
Break magnitude	M^*_{1450}	-25.13
Bright end slope	α	-2.81
Faint end slope	β	-1.5

where $\Phi(M_{1450}, \mathrm{z})$ is defined in units of $\mathrm{Mpc}^{-3}\,\mathrm{mag}^{-1}$ and is quoted with respect to the rest frame magnitude of the continuum at $\lambda = 1450\,\text{Å}$. This is a standard practise in the optical that allows for straight forward comparisons. Fitted parameters in equation 3.2.1 are the break magnitude M^*_{1450}, the normalisation constant at the break magnitude $\Phi(M^*_{1450})$ and the bright end power law index α. A detailed description of the fitting process can be found in the original publication and the parameter set adopted by Willott et al. is shown in table 3.1. A redshift dependent density evolution term k and the faint end power law index β are both unfitted and each is now afforded some further discussion.

It is well established that the space density of quasars peaks at $\mathrm{z} \simeq 2.5$ and declines rapidly thereafter [e.g. 108]. At $\mathrm{z} \simeq 6$, the space density of bright SDSS quasars ($M_{1450} \lesssim -27$) measured by Fan et al. [101] is consistent with a single power law extrapolation of the QLF between $3 \leq \mathrm{z} \leq 5$ [109]. The rate of evolution implied by these luminosity functions corresponds to a value $k = -0.47$, where the space density declines exponentially as $10^{k(\mathrm{z}-6)}$. Beyond $\mathrm{z} \simeq 6$ the value of k is of course completely unknown and moreover one cannot rule out the possibility of a noticeable luminosity dependence, similar to those inherent in quasar, AGN and galaxy populations at lower redshifts [110–112].

Clearly, by extrapolating much past $\mathrm{z} \simeq 6$ one should only make baseline projections. As noted by Willott et al. [43], some confidence can be taken from recent theoretical modelling which shows that for black holes accreting close to the Eddington limit the global black-hole accretion rate density increases by roughly 10 to 100 times between $\mathrm{z} = 9$ and $\mathrm{z} = 6$ [113, 114]. In this context, these results should be viewed with some caution since they focus on isolated regions of high density. Nevertheless, such an evolution broadly agrees with the decline in space density adopted here, which evolves by a factor 30 over the same interval.

The faint end slope β is also unfitted and is set based on measurements at lower redshifts [110, 115, 116]. If there is minimal evolution in the break magnitude (i.e. negligible luminosity dependence) beyond $\mathrm{z} = 6.0$, then VIKING will probe the bright end of the luminosity function and uncertainty in the faint end slope will have a small effect on number count projections. If evolution in the break magnitude to $\mathrm{z} = 7.5$ is considerable, then number count projections based on the Willott et al. baseline luminosity

function could depart from the true number significantly. With the above discussion in mind, one must interpret any inferences derived via extrapolation of the QLF to earlier times with necessary caution.

The number of quasars brighter than an apparent magnitude $m_{\rm i}$ in some fixed observing passband i over a given redshift interval may be calculated via integration of the QLF over comoving volume V_c to a depth given by the absolute magnitude limit $M_{1450,\rm lim}(\rm z)$. The comoving volume element is related to the angular diameter distance D_A, the Hubble parameter $H(\rm z)$ and the survey area Ω by

$$dV_c = \frac{c\,(1+\rm z)^2\,{D_A}^2}{H(\rm z)}\,d\Omega\,d\rm z, \tag{3.2.2}$$

where $D_A = (1+\rm z)^2 D_L$. For a flat universe the luminosity distance is given by

$$D_L(\rm z) = c\,(1+\rm z)\int_0^{\rm z}\frac{d\rm z}{H(\rm z)} \tag{3.2.3}$$

and $H(\rm z)$ is given by

$$H(\rm z) = H_0\,\sqrt{\Omega_M(1+\rm z)^3 + \Omega_k(1+\rm z)^2 + \Omega_\Lambda}, \tag{3.2.4}$$

where Ω_M, Ω_k and Ω_Λ are the energy densities in the matter, curvature and cosmological constant of the Universe. The final integration then takes the following form

$$N_q = \int_0^\Omega d\Omega\int_0^{\rm z}\int_{-\infty}^{M_{1450,\rm lim}(\rm z)}\frac{c\,{D_L}^2}{H(\rm z)(1+\rm z)^2}\Phi\,d\rm z\,dM_{1450}, \tag{3.2.5}$$

where the integral over solid angle Ω may be replaced by the survey area in steradians.

Since we do not measure the magnitude of a quasar at the rest frame wavelength, but rather in some bandpass fixed in the observers frame, one needs to first convert from the flux or equivalently magnitude measurement in the redshifted frame to the corresponding measurement in the rest frame. This is the K-correction, defined by

$$M_{1450} = m_{\rm i} - 5\log_{10}(D_L/10\,\rm pc) - K_{\rm corr}. \tag{3.2.6}$$

For a featureless quasar spectrum, the continuum is well modelled by a power law in frequency ν, such that $f_\nu \propto \nu^\gamma$, so the K-correction to rest frame wavelength $\lambda = 1450\,$Å from an arbitrary bandpass with central wavelength $\lambda_{\rm i}$ is

$$K_{\rm corr} = -2.5\,(1+\gamma)\log_{10}(1+\rm z) - 2.5\,\gamma\log_{10}(1450\,/\lambda_{\rm i}), \tag{3.2.7}$$

where γ is taken to be $\gamma = -0.5$, in accordance with the composite quasar spectrum of

Vanden Berk et al. [1] and continua of $z \sim 6$ SDSS quasars [e.g. 11, 101]. The VIKING quasar search will select objects based primarily on their magnitude in the J band. The central wavelength of VISTA's J band filter is $\lambda_J = 12500$ Å and so evaluating equation 3.2.5 numerically for the case where $k = -0.47$ and K-correcting from λ_J gives the projected baseline number count model shown in figure 3.5 over the interval $6.5 \leq z \leq 7.5$. For comparison the non-evolving $k = 0$ and highly-evolving $k = -0.94$ models are plotted alongside, as shown by the short and long dashed lines respectively. The latter corresponds to a factor of 3 decline in density per 0.5 increase in redshift.

The thick grey regions show the range of deviation from each curve when the faint end slope is allowed to vary between $-2 \leq \beta \leq -1$ and the break magnitude is kept constant. In this case, variation in the faint end slope has only a weak effect on number count projections. The best constraints on the break magnitude at lower redshifts show evolution in the range $-26 < M_{1450} < -24$ over the interval $1 < z < 4$ [108, 110], where several forms of joint and independent density and luminosity evolutions have been investigated. Each is entirely plausible beyond $z = 6$ and so to capture a feeling for the effect that an evolving break magnitude might have on quasar number count projections, it seems fairest to consider an average result by evolving M^* linearly in flux units with redshift in the range $-26 \leq M^* \leq -24$, whilst keeping the normalisation constant. The thin red and blue lines in figure 3.5 show the baseline model after applying evolution from $M^* = -25.13$ to $M^* = -24.00$ and $M^* = -26.00$ respectively. Clearly, significant departures from baseline evolution are possible with modest changes in evolutionary behaviour beyond $z = 6$. This highlights the large uncertainties in the calculation.

Naturally, one would like to select quasars as faint as possible, but experience from other surveys [e.g. 101] has shown that a $\sim 10\sigma$ detection in at least one band is necessary to avoid the selection of large numbers of false positives as a result of photometric scatter. Therefore, figure 3.5 also plots the inverse area of VIKING, the LAS and the VHS at their 10σ limiting J band magnitudes as filled circles. This metric can be equivalently thought of as the surface density at which a given survey would contain a mean of 1 quasar. It is clear from the figure that VIKING is predicted to contribute significantly to increasing numbers of near-infrared selected $z \gtrsim 6.5$ quasars, surpassing the numbers expected from the UKIDSS-LAS and providing a similar haul to that expected from the VHS. To summarise, table 3.2 gives the number of quasars predicted by each of the number count models at various limiting magnitudes in each survey. Note that expected numbers from the UKIDSS-LAS are given for data acquired up to DR8.

Lastly, it was encouraging to hear of recent results from the latest UKIDSS data release, DR8, which contains some 2600 deg^2 of LAS imaging in at least the Y and J bands. Referring to table 3.2, the baseline number count model predicts that there should be a mean of 1.91 ± 1.38 quasars in this area, while the non-evolving and highly-evolving models predict 5.12 ± 2.26 and 0.78 ± 0.88 respectively. At present there is a single

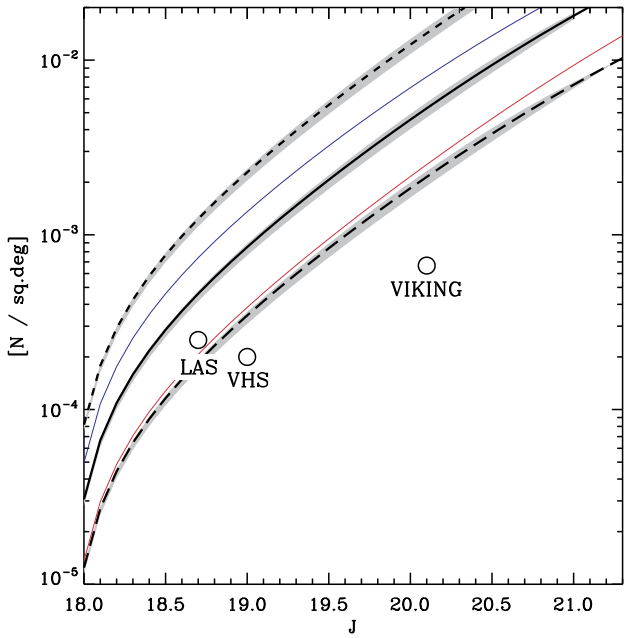


Figure 3.3: Number of quasars brighter than apparent J-band magnitude J, over the redshift interval $6.5 \leq z \leq 7.5$, for three models based on the luminosity function of Willott et al. [4]. The parameter k determines the rate of density evolution beyond z = 6. The short-dashed, solid and long-dashed curves correspond to $k = (0.0, -0.47, -0.94)$ respectively (see text for more details). Open circles have been placed at the inverse area of various surveys, showing the surface density at which the given survey would contain a mean of 1 quasar brighter than its nominal 10σ detection limit. Shaded regions show the deviations in each curve when the faint end slope ($\beta = -1.5$) is allowed to vary between $-2.0 \leq \beta \leq -1.0$ with a constant break magnitude. The thin red and blue curves show the effects of applying break magnitude evolution in the baseline model from $M^* = -25.13$ to $M^* = -24.00$ and $M^* = -26.00$ respectively.

Table 3.2: $6.5 \leq z \leq 7.5$ quasar number counts in VIKING, the LAS and the VHS, derived from the QLF of Willott et al. [4] for the pure density evolutionary models described in the text. Errors are 68% poissonian credible intervals (see Chapter 5) only since the uncertainties inherent in extrapolating the QLF from $z > 6$ are unknown, these are discussed qualitatively in the text. Note also that VIKING and VHS number count predictions are given for the full survey areas, whereas the LAS predictions are given up to DR8; $\sim 2600\,\mathrm{deg}^2$ in at the least Y and J bands.

VIKING

Depth	$k = 0.00$	$k = -0.47$	$k = -0.94$
5σ ($N < J = 20.9$)	$64.86^{+9.05}_{-7.05}$	$24.01^{+5.94}_{-3.95}$	$9.68^{+4.20}_{-2.21}$
8σ ($N < J = 20.4$)	$33.64^{+6.83}_{-4.83}$	$12.49^{+4.62}_{-2.60}$	$5.05^{+3.37}_{-1.38}$
10σ ($N < J = 20.1$)	$22.03^{+5.75}_{-3.74}$	$8.20^{+3.95}_{-1.97}$	$3.32^{+2.98}_{-0.99}$
20σ ($N < J = 19.4$)	$7.68^{+3.86}_{-1.88}$	$2.87^{+2.86}_{-0.88}$	$1.16^{+2.39}_{-0.34}$

LAS DR8

Depth	$k = 0.00$	$k = -0.47$	$k = -0.94$
5σ ($N < J = 19.6$)	$18.14^{+5.76}_{-3.31}$	$6.76^{+3.70}_{-1.72}$	$2.74^{+2.83}_{-0.84}$
8σ ($N < J = 19.1$)	$8.30^{+3.98}_{-1.97}$	$3.10^{+2.92}_{-0.93}$	$1.26^{+2.38}_{-0.38}$
10σ ($N < J = 18.8$)	$3.94^{+3.13}_{-1.14}$	$1.91^{+2.60}_{-0.60}$	< 1.0
20σ ($N < J = 18.1$)	$1.61^{+1.9}_{-0.80}$	< 1.0	< 1.0

VHS

Depth	$k = 0.00$	$k = -0.47$	$k = -0.94$
5σ ($N < J = 19.9$)	$54.82^{+8.39}_{-6.41}$	$20.41^{+5.57}_{-3.57}$	$8.27^{+3.96}_{-1.97}$
8σ ($N < J = 19.4$)	$25.61^{+6.09}_{-4.12}$	$9.55^{+4.17}_{-2.19}$	$3.88^{+3.12}_{-1.13}$
10σ ($N < J = 19.1$)	$15.96^{+5.04}_{-3.05}$	$5.96^{+3.55}_{-1.57}$	$2.42^{+2.74}_{-0.75}$
20σ ($N < J = 18.4$)	$5.12^{+3.39}_{-1.40}$	$1.91^{+3.55}_{-0.83}$	< 1.0

reported discovery of a $z > 6.5$ quasar in the LAS DR8 [$z = 7.085$ to be more precise; 37], which marginally favours evolving models.

3.2.2 Stellar number counts

The Besançon stellar population synthesis model of the Galaxy [BGM; 117, 118] allows stellar number counts up to spectral type M to be computed as function of J-band magnitude and direction on the sky via an interactive web form[6]. The model is currently undergoing an update to the set of M-dwarf atmospheres it uses. The current study therefore provides an ideal testbed in which to compare their performance. A full description of the model is given in the above cited work.

The atmospheres of cooler L- and T-dwarfs are much less well constrained and so a more empirical approach is taken to calculate their number counts.

The Besançon Stellar Population Synthesis Model of the Galaxy: M-dwarf number counts

The BGM models the Galactic thin-disc, thick-disc, halo and bulge via a specific treatment of the density laws, the initial mass function (IMF) and the star formation rate in each component.

At the low masses and faint magnitudes considered here, the scale height of the relevant component and the slope of the IMF at the low mass end are of importance. In the thin disc the scale height is constant for of an isothermal population of given age and vertical velocity dispersion.

The IMF used here differs from the on line version which employs a very steep slope in the thick disc [119]. Consequently the online version of the code over estimates the number of M-dwarfs when compared with deep data. The authors attribute this to an error in the default calibration of the model in the CFHTLS bands. The revised slope, used here down to very low masses is discussed by [120]. A value $\gamma = 0.7$ is used where the IMF is parametrised as $dN/d\mathrm{M} \propto \mathrm{M}^{-(1+\gamma)}$.

The BGM magnitudes are computed from a set of stellar atmosphere models. Most of the HR-diagram is covered by the Basel library [121], but these models depart significantly from observations at low masses [119] and thus, for effective temperatures below $\sim$ 4000K NextGen model atmospheres [122] are employed.

Stellar number counts were calculated on a 2D grid of spectral-type and J-mag bins for 30 representative fields of 10 deg^2 for each of the thin-disc, thick-disc and halo components. The results were scaled up accordingly. In order to make certain that density gradients were well sampled, effort was made to ensure that each field was offset from

[6] http://model.obs-besancon.fr/

adjacent fields by no more than 5 deg of Galactic latitude. J-band number counts were then calculated in the range $16.0 \leq \mathrm{J} \leq 24.0$ in bin widths of $\Delta\mathrm{J} = 0.5$. According to the quasar number count model this is $\sim$ 2 mag brighter than we can realistically expect to recover quasars in VIKING and $\sim$ 3 mag fainter than the 5σ limiting magnitude of the survey. The extra range is included to allow sources to scatter in and out of the survey as the result of photometric error.

L and T dwarf number counts

The main structural components of the Milky Way have been identified with star count data and accompanying colour information in a number of classic and modern works [e.g. 6, 123]. Gilmore [124] was the first to show that the Galaxy could be well represented by three distinct components; an old metal deficient halo population represented by a deprojected de Vaucouleurs profile and young and intermediate populations whose total contribution is given by the sum of two double exponential profiles describing a concentrated 'thin-disc' and a more rarefied 'thick-disc'. The space density of either disc component at Galactic coordinates (l, b) and heliocentric distance d assumes the following form

$$\rho_{disc}(d,l,b) \propto \exp\left[\frac{-R(d,l,b) - R_{\odot}}{L}\right] \exp\left[-\frac{|Z_{\odot} + d\sin b|}{H}\right], \tag{3.2.8}$$

where the proportionality constant is the spatial density of the relevant object class or classes at $Z = 0$ i.e. at the Galactic plane. $Z_{\odot}$ is the position of the sun relative to the Galactic plane, H and L represent the scale height and length of the particular disc component and $R_{\odot}$ and $R(d,l,b)$ are the galactocentric distances of the sun and the point of interest respectively. The latter of these is given by

$$R(d,l,b) = (R_{\odot}^2 + d^2\cos^2 b - 2\,R_{\odot}\,d\cos b\cos l)^{1/2} \tag{3.2.9}$$

which can be re-arranged to give

$$R(d,l,b) = [(R_{\odot} - d\cos b\cos l)^2 - d^2\cos^2 b\cos^2 l + d^2\cos^2 b]^{1/2}. \tag{3.2.10}$$

and since VIKING will observe extra-galactic fields i.e. out of the Galactic plane, the assumption $R_{\odot} \gg d$ is valid and equation 3.2.9 may be approximated as

$$R\,(d,l,b) \approx R_{\odot} - d\cos b\cos l. \tag{3.2.11}$$

In substituting this back into 3.2.8 we are left with

Table 3.3: Parameters of the Galactic thin and thick disc system [5, 6].

$R_\odot$ (pc)	$Z_\odot$ (pc)		L (pc)	H (pc)
8600 ± 200	$+27 \pm 4$	Thin:	2250 ± 1000	330 ± 3
		Thick:	3600 ± 720	580 - 750
		Thick disc normalisation: 13 - 6.5 %		

Table 3.4: The local space density $\rho_\odot$ and absolute J-band magnitude M_J of field L and T dwarfs from Caballero et al. [7].

	L0-1	L1-2	L2-3	L3-4	L4-5	L5-6	L6-7	L7-8	L8-9
M_J	11.72	12.19	12.57	12.95	13.33	13.71	14.01	14.40	14.68
$\rho_\odot$	0.636	0.861	1.000	0.834	0.819	0.869	0.785	0.644	0.462
	L9-T0	T0-1	T1-2	T2-3	T3-4	T4-5	T5-6	T6-7	T7-8
M_J	14.45	14.38	14.10	13.85	13.77	14.21	14.85	15.10	16.1
$\rho_\odot$	0.308	0.222	0.241	0.406	0.788	1.42	2.38	3.45	4.82

$$\rho_{disc}(d,l,b) \propto \exp\left(\frac{d\cos b\cos l}{L}\right)\exp\left(-\frac{|Z_\odot + d\sin b|}{H}\right), \tag{3.2.12}$$

which can be evaluated numerically (or analytically) for any direction over which our assumptions are valid.

The local space densities of cool stars of spectral type L0 to T8 are given as a function of their absolute J-band magnitudes in table 3.4 [taken from 7]. For consistency number counts were calculated for each spectral type in J-magnitude bins of width 0.5 over the range $16 \leq J \leq 24$ and in the same fields as simulated by the the Besançon model for M-dwarfs. The distance d of any object with apparent magnitude J is determined via the distance modulus, $\vartheta = J - M_J$ by $d = 10^{0.2\,\vartheta+1}$ and the parameters of the thin and thick discs listed in table 3.3 are taken from [5] and [6].

An identical luminosity function in both the thin and thick discs is assumed. This probably overestimates the count of thick-disk brown dwarfs, since most of these objects are predicted to have very faint magnitudes by theoretical cooling sequences [105]. This would place most of the thick-disk population with $M < 0.06M_\odot$ below VIKING detection limits. This assumption is therefore conservative since it will overestimate the brown-dwarf contamination of the quasar sample.

A similar argument can be made for the halo, which represents an even older population. In this case we assume that all halo L and T-dwarfs will be below VIKING detection limits. BGM number counts of halo M-dwarfs fall off rapidly towards cooler spectral types. The BGM surface density of halo M9-dwarfs with $J < 23.0$ is $< 1\,\mathrm{deg}^{-2}$, suggesting that this is a reasonable approach.

3.2.3 Catalogue generation

With stellar number counts binned by apparent J-band magnitude and spectral type the next step is to pair each count with a set of synthetic photometry to mimic an observation of each source.

For N_c sources distributed over N_{i-1} magnitude and N_{j-1} spectral type intervals, the procedure is to step over all $N_i \times N_j$ bins iteratively. Each object receives a set of synthetic fluxes, $\mathbf{F}_c = \{f_Z, f_Y, f_J\}$. $\mathbf{F}_c$ is generated via interpolation of our synthetic photometry (section 3.1) between a source of the relevant spectral type T_j and that of the later adjacent spectral type, T_{j+1}. If the k^{th} occupant of bin $i \times j$, has a magnitude J in the range $J_i \leq J < J_{i+1}$ and a spectral type T_j, then $\mathbf{F}_c$ is defined by the following weighted sum,

$$\mathbf{F}_c = \varepsilon\, \mathbf{F}_{i,j,k} + (1-\varepsilon)\, \mathbf{F}_{i,j+1,k}, \tag{3.2.13}$$

where ε is a uniform random deviate in the range $0 < \varepsilon \leq 1$. Each of the components of $\mathbf{F}_c$ (i.e. $\mathbf{F}_{i,j,k}$ and $\mathbf{F}_{i,j+1,k}$), comprise of a Z, Y and J band flux generated by drawing Z and Y band fluxes from our library of synthetic photometry for the relevant spectral type. These are then scaled to a reference flux f_J, which is derived from a magnitude J drawn randomly from the interval $J_i \leq J < J_{i+1}$. This process is repeated over all $N_i \times N_j$ bins until all number counts are coupled to a set of unique synthetic photometry.

The VSA VIKING release v1.0 with data up to July 2010 contains 173 image tiles. During the image reduction and pipeline processing the average sky noise in each image tile is measured by placing a number of circular apertures in sparsely populated regions. These measurements are then placed in the header units of each image before being uploaded to the VSA. In a real survey of faint sources the distribution of measured flux errors is sky dominated and usually well approximated as being Gaussian. The intrinsic distribution of synthetic fluxes can therefore be perturbed by drawing from a set of Gaussian distributions, with standard deviations equal to the sky noise level in the VIKING image tiles and scaled to the theoretical flux of the object in the corresponding passband. Adding this distribution of errors to the set of simulated intrinsic fluxes gives a set of simulated 'observations'.

The catalogue extraction process for the VIKING survey, is handled by the VDFS at CASU. The CASU source extraction algorithm employs a detection threshold of $\sim 5\sigma$ but this is not a rigid limit. Thus to determine whether our synthetic sources would be detected by the CASU source extractor it is necessary to model the signal-to-noise distribution of detected sources from the data. Here we follow a novel method used by Patel [125], which is described in brief below.

Since we are modelling a population of point sources it is necessary to limit the data analysis to stars only, which requires a highly reliable method of star-galaxy separation.

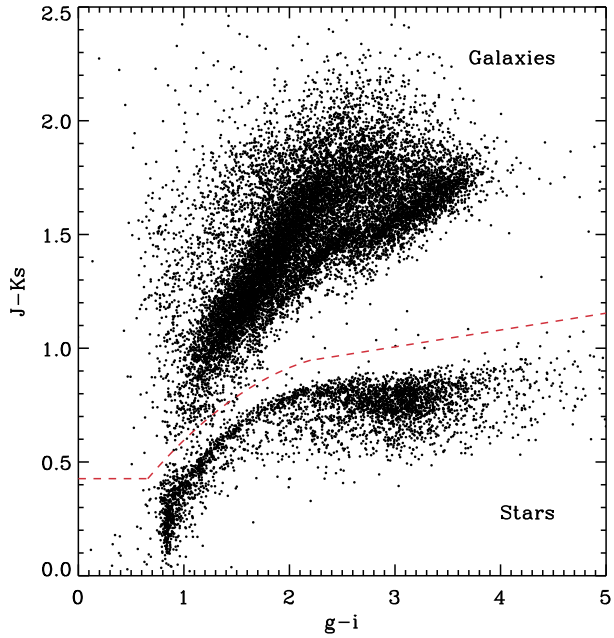


Figure 3.4: The CFHTLS-VIDEO *gi*JKs colour plane for matched VIKING-VIDEO-CFHTLS sources in the $\sim 1\,\mathrm{deg}^2$ region of overlap shared by each survey. The deep CFHTLS-VIDEO observations allow robust star galaxy separation. The red dashed line shows the regions in which the star and galaxy populations are defined.

In contrast to Patel [125], who used morphological classifiers to define a reliable and complete sample of point-like objects. Here the VVC catalogue, compiled in chapter 1, is used to separate stars and galaxies via their optical-near-infrared colours. At the time of writing only VIDEO Y JHKs observations are available, but the lack of Z-band data is not important since most high-z quasars will be too faint to be detected in VIKING Z-band observations given the CASU $\sim 5\sigma$ extraction criteria.

The VIDEO-CFHTLS $g-i$ vs. J − Ks plane is shown in figure 3.4. The extra depth in the matched catalogue minimises photometric scatter and with the combination of colours gives remarkably good star galaxy separation. The red dashed line separates two regions from which are drawn a sample of stars and galaxies. Where possible i or g dropouts have been classified by virtue of either their z − J vs J − Ks colours or their J − Ks colours alone. We will look at the galaxy sample in later sections, this section focuses on the stellar sample.

Histograms of the Y and J-band magnitude distributions of matched stars are shown in figure 3.5 for VIKING (red) and VIDEO (blue) photometry. For comparison, the full distribution of VIDEO-CFHTLS matched point sources is also plotted (black histogram). Synthetic sources were drawn randomly from this distribution and each source was perturbed by a Gaussian error with a standard deviation equal to the sky-noise level drawn from VIKING image tiles. Each perturbed source was then given a detection

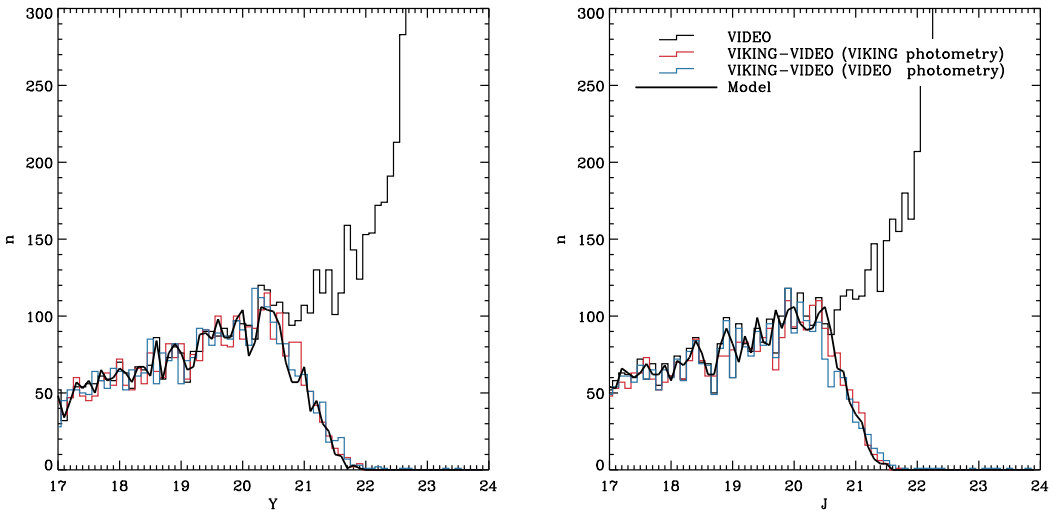


Figure 3.5: The Y and J-band magnitude distributions of sources matched between VIDEO-CFHTLS (black histogram format) and VIKING-VIDEO-CFHTLS with VIKING photometry (red histogram format) and with VIDEO photometry (blue histogram format). The black continuous curve shows the distribution of simulated sources with intrinsic magnitudes drawn from the VIDEO-CFHTLS matched catalogue and perturbed with realistic VIKING-like Gaussian errors. A Gaussian distributed signal to noise cut has been applied to mimic the source extraction process used to produce the VIKING catalogue (see text for details).

significance drawn from a Gaussian distribution with mean μ_i and standard deviation σ_i, where the subscript refers to the arbitrary bandpass i. If the source passed the simulated detection cut-off then it was kept otherwise it was discarded. The resulting distribution was then compared by eye to the VVC matched distribution and the process was repeated with various values of μ_i and σ_i until a good match was found. The continuous black curve in figure 3.5 shows the best matched distribution of simulated sources. This was produced using a Gaussian distribution of simulated detection significances with $\mu_J = 0.22$ and $\sigma_J = 0.04$. The same procedure was undertaken in the Y-band yielding a best match Gaussian with $\mu_Y = 0.15$ and $\sigma_Y = 0.06$. The results were applied to the set of simulated observations described above resulting in a realisation of the cool star component of the VIKING catalogue over the entire 1500 deg^2 of the VIKING field.

3.3 Colour space contamination

The magnitude and colour distribution of the simulated stellar population has been compared with data from the VVC catalogue. The results are shown in figure 3.6, where the data is plot in histogram format and the simulated catalogue as smooth curves. Both the magnitude and colour distributions of the model are in overall good

agreement with the data.

An excess of blue stars can be seen in the colour distributions in the left hand panels. Closer inspection of the magnitude distribution of these stars shows that they are reasonably faint with number counts peaking at $Z \sim 19.5$, $Y \sim 19.5$ and $J \sim 19.0$. Given their blue colours, they are likely intrinsically bright main sequence halo stars. To promote a fair comparison between the true and modelled magnitude distributions an effort has been made to remove the majority of the blue stars from the data by retaining only those having $Z - Y \geq 0.0$, $Y - J \geq 0.0$, $Z - J \geq 0.3$.

The overall agreement between colour and magnitude distributions in the model and the data is good, but the model over predicts the number counts in some bins by up to a factor of 1.3. The number count predictions are therefore conservative. In the modelled population almost all objects are M-dwarfs whose number counts and J-band magnitudes are predicted by the BGM. Given the rarity of cooler type objects, the 1 deg^2 patch of deep VVC matched imaging is not sufficient to constrain the L- and T-dwarf number count model well. The overall good agreement with the data suggests that for the latest type dwarfs, the modelled objects are distributed accurately over colour and magnitude space and that the predicted number counts are within an order of magnitude of their true values. It is also encouraging that the results broadly conform with data from recent near-infrared searches [88, 89].

3.3.1 Stellar contamination

The full VIKING cool-star mock catalogue (i.e. 1500 deg^2) is plotted in ZYJ colour space in figure 3.7. Contours enclose bins of width $\Delta ZYJ = 0.1$ with occupancies decreasing logarithmically from the centre of the contour map outwards as 10^5, 10^4, 10^3 and 10^2. The filled black circles denote individual objects belonging to bins with occupancies of less than 10^2. Also shown is a catalogue of ~ 200 simulated $6.5 \leq z \leq 7.5$ quasars. Each source has been 'detected' at the 8σ level in Y and J and was drawn from a larger catalogue of ~ 1000 quasars with a uniform distribution in redshift.

Over plotted in red is the selection region described by

$$\begin{aligned} Z - Y &\geq 1.2 \\ Y - J &\leq 1.0, \end{aligned} \tag{3.3.1}$$

which was shown to reject the majority of Galactic M, L and T dwarfs and accept the majority of quasars in figure 3.2.

With the introduction of error perturbed photometry and accurate number count models, the accuracy of this statement has not changed, most stars are still rejected by this

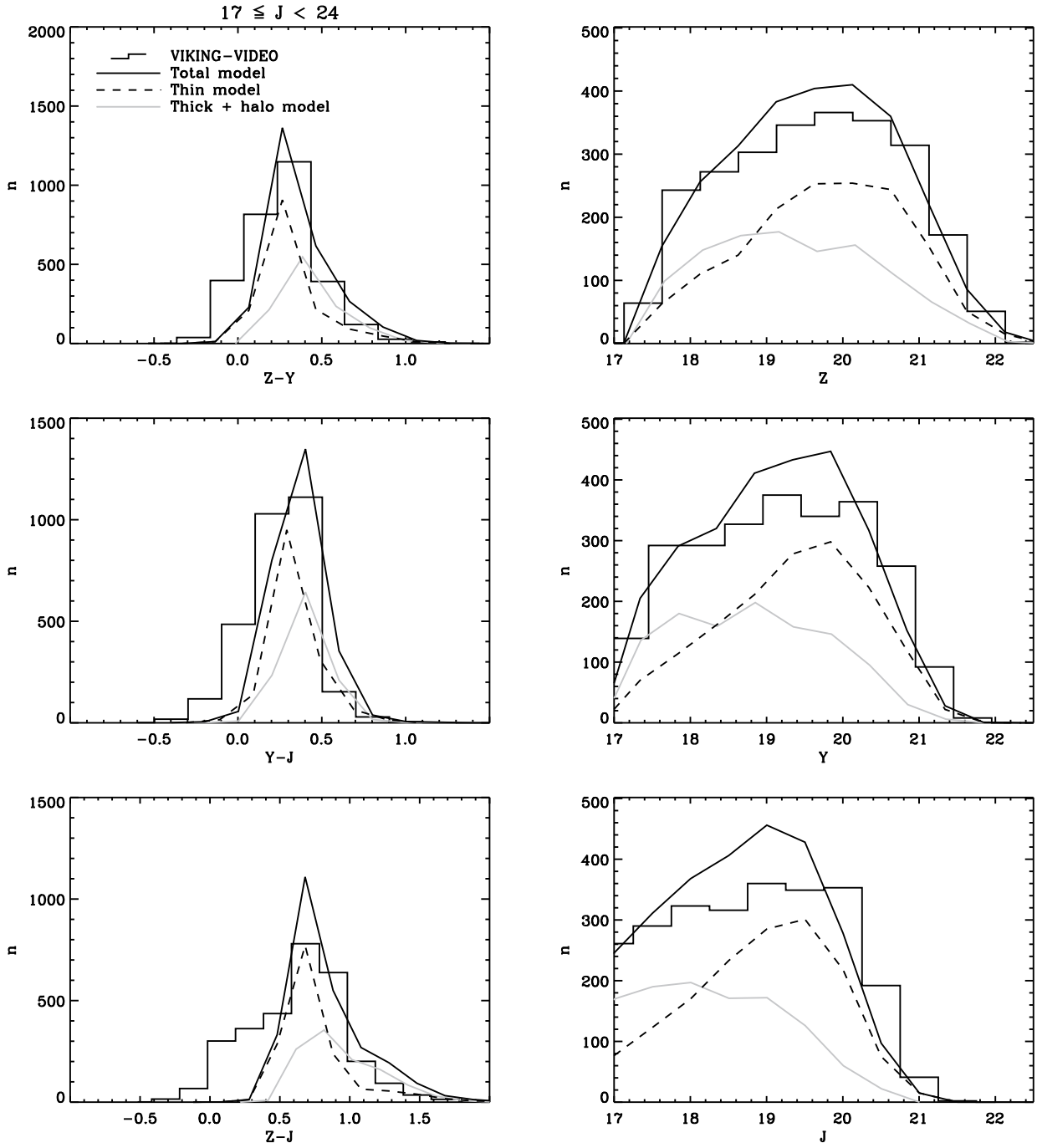


Figure 3.6: Left: VIKING Z, Y, J colour distributions of the simulated thin disc, combined thick disc and halo and total cool star contributions in $\sim 1\,\mathrm{deg}^2$ of the VIKING-VIDEO-CFHTLS overlap region. Each distribution is compared to the real distribution of VIKING stellar photometry as shown by the histogram style plots. The excess of blue objects in each plot is attributed to main-sequence halo stars with bright absolute magnitudes, an effort has been made to remove these objects in order to compare modelled and true magnitude distributions. Right: As in the left hand panel but for the simulated and real Z, Y, J magnitude distributions. In both cases the the overall agreement with the data is good. There is a slight overshoot in predicted number counts making these models conservative.

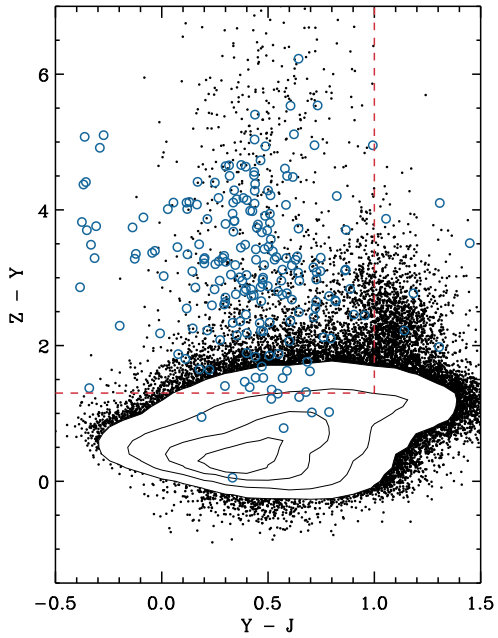


Figure 3.7: The VIKING ZYJ colour plane for simulated cool stars (contours and points). Also shown are $\sim$ 200 simulated quasars (open circles). The red dashed line shows the colour selection region defined by equation 3.3.1.

selection region while most quasars are retained, but the introduction of the model now shows that within these constraints stars will outnumber quasars by a factor $\sim 10^4$. For the classic rare object search in which candidates are chosen predominantly by virtue of their location in multi-colour spaces with perhaps detection significance, morphology etc. acting as qualitative secondary constraints, searching for $\sim$ 10 objects among these candidates would present a challenging task. What's more is that VIKING is a galaxy dominated survey so the inevitable contamination from extra-galactic objects will exacerbate the problem.

3.3.2 Extra Galactic contamination

Up to now, the focus of this chapter has been on contamination from foreground Galactic M, L and T-dwarfs. Referring back to figure 3.2 it is clear that contamination from z $\simeq$ 1 - 2 galaxies could also pose a problem. As we shall see in the next chapter, most well detected galaxies can be rejected on the basis of their extended disc-bulge morphologies. For completeness it is interesting at this point to look briefly at the population of galaxies which contaminate quasar selection space. The star galaxy separation technique discussed in section 3.2.3 affords us this opportunity.

The galaxy sample considered here, is on the whole slightly less reliable than the stellar sample considered in section 3.2.3, since it shares a similar region of colour space with low-z ($0.1 \leq z \leq 2.5$) quasars. The lack of any SDSS or other spectroscopic overlap

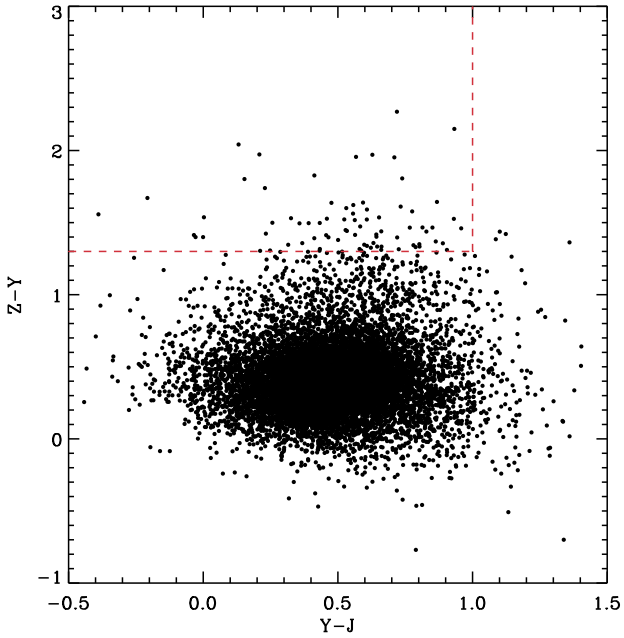


Figure 3.8: The VIKING ZYJ colour plane for a highly complete sample of galaxies selected from the VIKING-VIDEO-CFHTLS overlap region. The red dashed line shows the region defined by equation 3.3.1.

prevents any attempts to remove these interlopers with any confidence, but work on optical and near-infrared number counts [e.g. 108, 126] suggests that low-z quasars will be outnumbered by galaxies in this sample by at least two orders of magnitude and they can be reasonably neglected.

Figure 3.8 shows all colour selected galaxies in the VVC overlap region. Again the selection region described by equation 3.3.1 is over plotted in red. On large scales galaxies have an isotropic distribution on the sky. In $\sim 1\,\mathrm{deg}^2$ there will be some level clustering and cosmic variance will be an issue. Nevertheless, linearly scaling galactic number counts in this $\sim 1\,\mathrm{deg}^2$ region to the full $1500\,\mathrm{deg}^2$ of the VIKING field will give a broad idea of the contamination expected from extra-galactic sources. In doing this, contaminating galaxies are expected to outnumber quasars by $\sim 10^5$. Again this is an unacceptable level of contamination for a rare object search on any reasonable timescale.

3.4 Summary

The contamination of high-z quasar selection spaces by galaxies and cool stars is a well known and documented problem. Indeed most authors begin their quasar searches with an analysis of the contamination similar to that presented in figure 3.2 [72]. Invariably the selection cuts they lay down turn out to be too optimistic and most candidates are either not observed or progressively, they fall out of candidate lists as cuts

are revised with the acquisition of larger data sets that allow predictions about the contamination with regard to the error perturbed photometry and distribution on the sky. Section 3.3 is therefore simply a reminder of this and places the extent of contamination on a more quantitative footing.

The simulation of a realistic VIKING catalogue, allows constraints to be placed on levels of contamination in high-z quasar searches in VIKING. At the time of writing this chapter, there were $< 200\,\mathrm{deg}^2$ of proprietary data catalogued in the VSA. The first photometric followup of VIKING quasar candidates was scheduled for June 2011. In advance of this, quasar candidates were to be chosen based on selection constraints determined from current limited available data. The simulated catalogue therefore adds a complementary data set from which to test and constrain selection strategies. This is the subject of the next chapter.

The methods employed here may be useful in new and forthcoming surveys, if there is need to define selection strategies in anticipation of a large data set becoming available.

CHAPTER 4

Quasar candidate selection

The first science ready data product, a paw-print based catalogue release with coverage of $\sim 150\,\mathrm{deg}^2$, became available to the VIKING working group in April 2010. With limited area and a depth some ~ 0.4 magnitudes brighter than that expected from tile based products, the chance of a high-z quasar discovery from such a data set is modest.

Based on information supplied by CASU regarding the rate of data acquisition and the quantity of existing data yet to be made available, it was estimated that $\sim 300\,\mathrm{deg}^2$ of science ready products would be available by Summer 2011. Quasar number count models (chapter 3) predict ~ 2 quasars in the redshift range $6.5 \leq z \leq 7.5$ over an area of $300\,\mathrm{deg}^2$ brighter than the J-band 10σ point source sensitivity limit, which would make the discovery of a quasar a reasonable prospect. A proposal for follow-up imaging time was submitted to ESO in September 2010 and time was awarded over three nights in June 2011 under program ID 087.A-0655(A). The task in the months running up to June was to develop a strategy with which to select high-z quasars from the VIKING data.

Searches for high-z quasars in wide field imaging surveys generally take the form of optimisation problems. One aims to maximise the number of quasars successfully detected in the most efficient way possible. This is parametrised by the survey contamination, defined as the ratio of the number of selected quasars to the number of selected contaminants, and the survey completeness, defined as the ratio of the number of selected quasars to the total number of quasars within the surveyed area and magnitude limits.

The design of a high-z quasar selection scheme that is both efficient and complete necessitates an accurate knowledge of the properties of both the targeted objects and the contaminating population measured by the relevant survey. In chapter 3, models of luminosity, colour and spatial distributions of the high-z quasar and cool-star components of the VIKING catalogue were developed. In this chapter these models are used along with a sample of VIKING galaxies to develop and compare two different approaches to quasar selection. The basic principles of these strategies are encapsulated

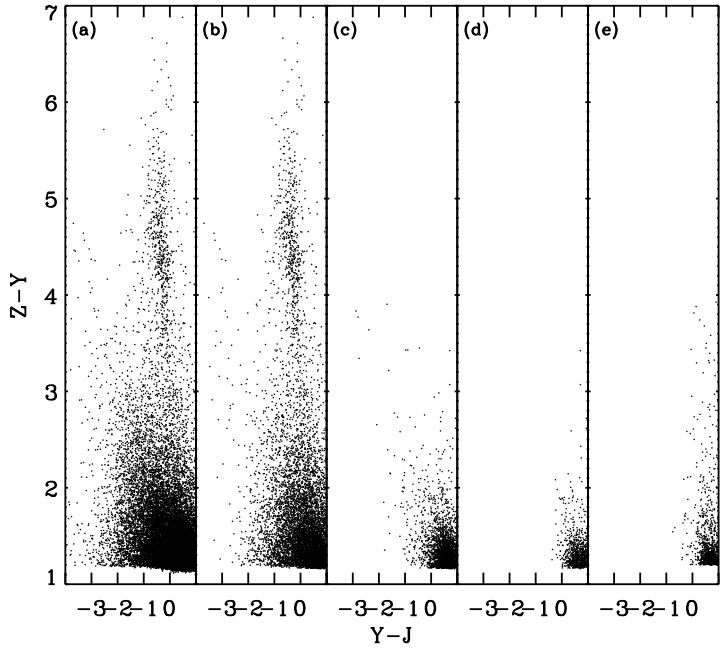


Figure 4.1: Increasingly exclusive constraints on the quasar candidate list isolate the most interesting objects. Panel (a) shows all candidates passing the broad colour cuts described by equation 3.3.1. Panel (b) shows remaining candidates after application of detection significance constraints in section 4.0.1. Panel (c) shows the catalogue after cuts on general low-quality detections, bright and saturated sources in section 4.0.2. Panel (d) shows remaining candidates after the series of morphological constraints are applied in section 4.0.3. Finally, panel (e) plots candidates after some further general constraints are applied and Z-band drop-outs have been re-measured as described in section 4.0.4.

by the traditional multi-colour selection techniques which have proven successful in the last decade [68, 72, 101] and the application of probabilistic candidate ranking, which has become popular in the last few years [127–130]. In both cases the search for high-z quasars begins in the VSA [82].

To begin the quasar search an obvious initial starting point, is with the set of colour cuts shown in figure 3.2 of chapter 3. A basic query for such objects in the VSA returns 3.18×10^5 candidates shown in ZYJ colour space in Figure 4.1 (a). This chapter describes the steps taken to reduce this candidate list, retaining all but the most promising objects for follow-up observation.

4.0.1 Detection significance

Any photometric measurement is made with a contribution from the sky background. For accurate measurement of the source brightness an estimate of the sky background flux must be subtracted from the total measured flux. For very faint sources the sky flux is dominant over the source flux and uncertainties in the sky background measurement

become the dominant source of error in the source flux measurement.

The effect of this uncertainty on the quasar search is that the fainter one decides to try and search for quasars the more uncertain the photometry becomes. The problem can be visualised in figure 4.1 (a), where this uncertainty brings significant scatter about the true distributions of stars and galaxies first considered in figure 3.2. It is desirable to select quasars as faint as possible, but experience from other surveys has shown that an 8σ to 10σ detection above the noise in the sky background in at least one band is necessary to avoid the selection of large numbers of false positives [e.g. 101]. High-z quasar candidates are necessarily faint in Z so a nominal photometric cut rejecting all candidates with $S/N \leq 8$ in both Y and J is adopted. The resulting candidate list contains the 6.8×10^4 objects shown is figure 4.1 (b), a factor 4.7 improvement.

4.0.2 General quality control

Some basic quality control reduces the catalogue by a further significant factor. Quality bit flags provided by the VSA encode quality issues pertaining to each detection. The relevant attributes are labelled *ppErrBits* in the VSA detection tables and are pre-fixed with the relevant filter name in the source tables. Bit flag values start at zero for detections with no known issues and increase towards issues of higher severity with up to 32 distinct values grouped into four different warning classes. The warning classes arrange issues as information, warnings, important warnings and severe warnings, with threshold bit flag values of 256, 65536 and 16777216 respectively. A detection whose *ppErrBits* are not within the information class has inaccurate photometry, which cannot be improved via direct re-measurement of the VIKING images and consequently, such sources are rejected as quasar candidates. This reduces the catalogue to 4.9×10^4 objects which includes those detections without quality issues, those that have been de-blended and those that include bad pixels.

Another quality issue arises from saturated sources, which have incorrect photometry. Quasar number count models, described in chapter 3 predict that $z \gtrsim 6.5$ quasars with $J < 18$ are unlikely to be found in the VIKING survey volume. Considerable uncertainties in the QLF warrant a cautious approach and hence objects brighter than the 17th magnitude in J are rejected. Further rejecting any object with $Y < 16$ ensures that none of the remaining candidates are saturated and reduces the total content of the catalogue to 4.4×10^4 sources, figure 4.1 (c).

4.0.3 Morphological constraints

The major focus of chapter 3, was to characterise the nature of cool-star contamination in VIKING quasar searches. This was appropriate since quasar searches in the SDSS, CFHQS and UKIDSS have all shown cool stars to be the most prominent of contami-

nants. The VIKING data set will differ from those of past wide field survey searches because the combination of depth and area will deliver a galaxy dominated data set. Galaxies in the range $1.2 \lesssim z \lesssim 1.4$ lie close the colour selection region shown in figure 3.2 and broadening of the galaxy locus by photometric scatter will place a fraction of galaxies in the selection box. Thus, a means to accurately identify galaxies would provide the potential to reduce the size of the candidate catalogue significantly.

Fortunately most galaxies can be rejected on the basis of their extended disc-bulge morphologies. At faint magnitudes, it is possible for the galactic disc component to fall below the detection limit of the survey, while the bulge component remains bright. Galaxies detected in this way can be morphologically indistinguishable from stars and usually the best way to avoid this type of contamination is to apply a detection significance constraint to ensure that candidates are sufficiently well detected in at least one band.

The star galaxy separation techniques discussed in chapter 3 present the opportunity to study the statistical properties of VIKING galaxies in detail and place further selection constraints on both the morphology and detection significance of quasar colour space contaminants.

The VSA provides a number of morphological classifiers based on the curve of growth of the independent detections of each source in a set of progressively inclusive apertures. The finer details of this approach are discussed extensively by Irwin et al. [131, in preparation], the main points are summarised as follows; measurements are made on ellipticity and curve-of-growth statistics. The results are then compared at the detector level to the average stellar locus as a function of magnitude. In this way each detection is given a morphological classification statistic known in the VSA as the *mergedClassStat* (*MCS*). For stars, this statistic is well approximated by an $\mathrm{N}(0,1)$ Gaussian distribution. Each individual classification is combined to give the overall classification for each source.

At bright magnitudes stars and galaxies populate two separate and well defined *MCS* loci. At faint magnitudes the galactic locus becomes indistinguishable from the stellar locus. This is demonstrated in figure 4.2 where the *MCS* is shown as a function of J for stars and galaxies selected from the VVC catalogue. The galactic locus begins to join the stellar locus at $\mathrm{J} \simeq 19$.

To investigate the best use of the *MCS* in discriminating against non-stellar morphology, a sample of objects which are unmatched between VIKING and VIDEO was added to the VVC catalogue and flagged as 'noise'. Figure 4.3 shows the completeness and contamination of the stellar sample by galaxies and noise within the following constraints; $-10 \leq MCS \leq 10$ (solid curve), $-7 \leq MCS \leq 7$ (dotted curve), $-5 \leq MCS \leq 5$ (dashed curve) and $-3.5 \leq MCS \leq 3.5$ (dash-dotted curve). It is clear that losses in completeness are small compared to gains in minimising contamination. However

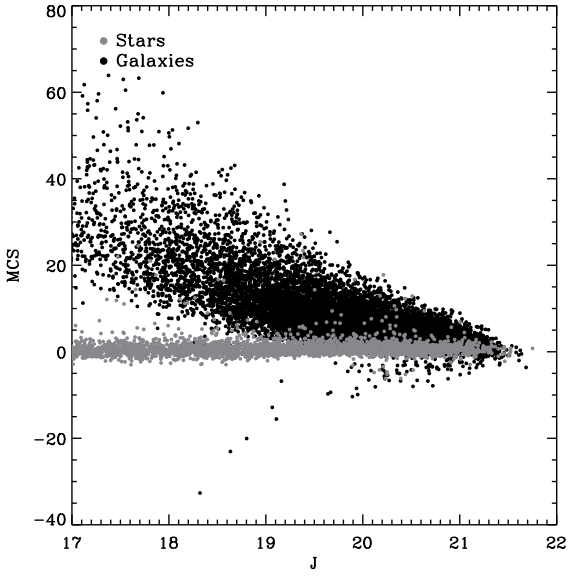


Figure 4.2: The *mergedClassStat* is plotted as function of J for stars and galaxies in grey and black respectively. The two populations occupy distinct loci and are separable at bright magnitudes. Increasing numbers of galaxies become indistinguishable from stars beyond J $\geq$ 19.

even with the tightest constraints contamination contributes to $\sim$ 50 per cent of the sample at the faintest magnitudes.

The VSA supplies a second useful set of morphological attributes called the *pClass* statistics, which assign four probabilities to each source: *pStar*, *pGalaxy*, *pNoise* and *pSaturated*. The *pClass* statistics are derived from a discrete classification flag based on the passband level *MCS* and several catches to make the classification more reliable (e.g. for saturated sources). Each discrete classification corresponds to a reasonably accurate, self-consistent probability for each of the allowed cases; star, galaxy, noise and saturated. The possible classification flags and associated probabilities are reproduced from the VSA in table 4.1.

Once each source has been given a passband level classification a combined, band merged, classification is given as follows,

$$pClass = \frac{\prod_{\mathrm{i}} P(class_{\mathrm{k}})_{\mathrm{i}}}{\sum_{\mathrm{k}} \prod_{\mathrm{i}} P(class_{\mathrm{k}})_{\mathrm{i}}} \tag{4.0.1}$$

where $class_{\mathrm{k}}$ is either star, galaxy, noise or saturated and i gives the ith band in which the source was detected. For example given an object detected in J,H and Ks and classified as -1,-2 and 1 respectively, using Table 4.1 and equation 4.0.1 one will find $pStar$ = 73.5%, $pGalaxy$ = 26.2%, $pNoise$ = 0.3% and $pSaturated$ = 0.0%. Note that since the *pClass* statistics are derived from a set of discrete probabilities, they too may only take on certain discretised values.

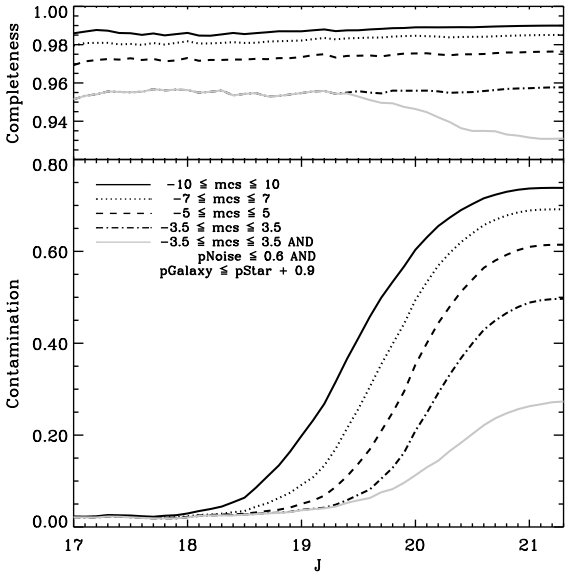


Figure 4.3: Upper: The completeness of the VIKING-VIDEO-CFHTLS stellar sample with various morphological selection constraints imposed (see legend and main text). Lower: The contamination of the VIKING-VIDEO-CFHTLS stellar sample by galaxies and noise with various morphological selection constraints imposed. All morphological statistics are derived from VIKING imaging.

Flag	Meaning	Probability			
		Star	Galaxy	Noise	Saturated
-9	Saturated	0.0	0.0	5.0	95.0
-3	Probable galaxy	25.0	70.0	5.0	0.0
-2	Probable star	70.0	25.0	5.0	0.0
-1	Star	90.0	5.0	5.0	0.0
0	Noise	5.0	5.0	90.0	0.0
+1	Galaxy	5.0	90.0	5.0	0.0

Table 4.1: Discrete morphological classification flags used by the VSA to assign the *pClass* statistics. Every detection in the VSA is flagged, with each flag corresponding to a reasonably accurate self consistent probability that the source is either a star, galaxy, noise-like or saturated. These probabilities are combined across passbands using equation 4.0.1 giving the *pClass* statistics.

The VVC catalogue is restricted to sources fainter than J $= 17.0$ and so in all cases $pSaturated = 0.0$ and $pStar + pGalaxy + pNoise = 1.0$. Figure 4.4 shows the VVC catalogue in *pClass* space. *pGalaxy* and *pStar* are plotted on the x- and y-axes respectively, while lines of constant *pNoise* are given along the diagonals. The top left hand panel plots colour selected stars. The number of objects at each point is indicated by the logarithmic colour bar. Similar plots are shown for colour selected galaxies and noise (i.e. VIKING objects without a counterpart in VIDEO) as labelled. The discretised nature of the *pClass* statistics is evident from the plots.

A large fraction of the noise and the majority of galaxies can be rejected with following morphological cuts,

$$pNoise \leq 0.6 \tag{4.0.2}$$

$$pGalaxy \leq pStar + 0.9 \tag{4.0.3}$$

which are shown by the shaded regions in the figure 4.4

These constraints are placed on the VVC catalogue and in figure 4.3 the contamination and completeness of the stellar component from noise and galaxies with $-3.5 \leq MCS \leq 3.5$ is shown in grey. With the new constraints the contamination from galaxies is significantly reduced with only a small reduction in completeness as the penalty.

Figure 4.5 (a) plots all VVC galaxies in the ZYJ plane. Galaxies passing the morphological constraints are plotted in panel (b), these galaxies are all indistinguishable from stars. Panel (c) shows the galaxy sample after the morphological, quality control and detection significance constraints have been applied. Only a small fraction of the original galaxies remain. With the introduction of the morphological constraints 1.1×10^4 quasar-like objects remain in the candidate catalogue. These objects are shown in figure 4.1 (d).

4.0.4 General constraints

A small fraction of the remaining objects can be rejected via a few more considered 'broad-brush' constraints.

VIKING shares a region of partial overlap with the SDSS foot-print in the ngc. At long wavelengths throughput over the SDSS z-band is dominated by the quantum efficiency of the SDSS charged coupled devices (CCD), which become unresponsive at $\lambda \simeq 9000$ Å. Consequently quasars with z > 6.5 should be dark to the optically sensitive SDSS instrument and any object detected by the SDSS in the VIKING-SDSS overlap region can be rejected as a high-z quasar candidate. The VSA tables are cross-matched with the SDSS DR7 and since the SDSS tables are held locally they can be accessed

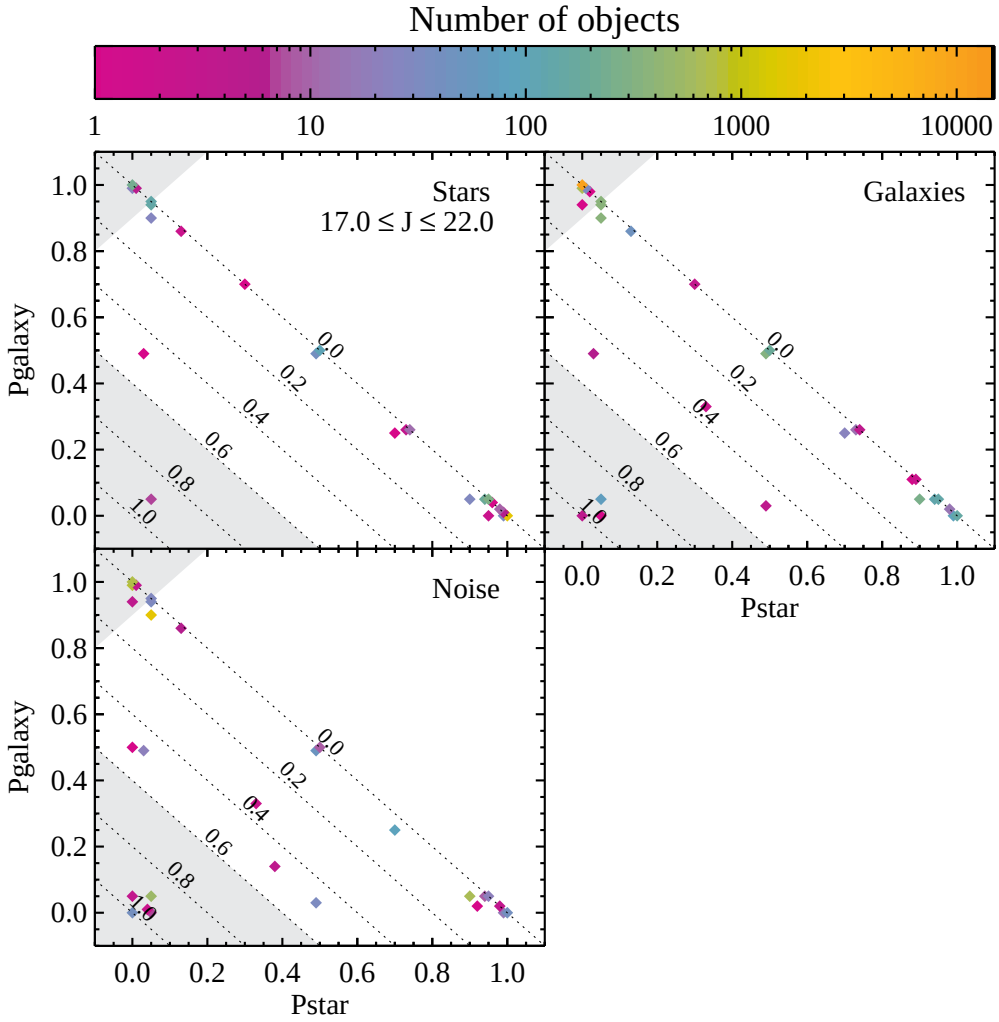


Figure 4.4: The VVC matched catalogue plotted in *pClass* space for colour selected stars in the top left, colour selected galaxies in the top right and objects labelled noise (i.e. detected in VIKING with no counterpart in VIDEO) in the bottom left. The *x*- and *y*-axes plot *pGalaxy* and *pStar* respectively, while lines of constant *pNoise* are plotted on the diagonals. The number of objects at a given point in *pClass* space is indicated by the colour bar. Shaded regions show the morphological cuts (equations 4.0.2 and 4.0.3) placed on the VVC catalogue in order to remove galaxies and noise without also removing significant numbers of stars (see main text for more details).

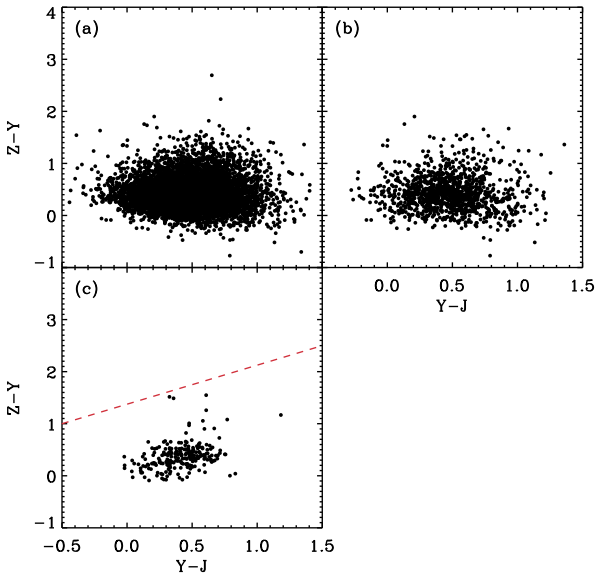


Figure 4.5: The VIKING Z, Y, J colour plane for a highly complete sample of galaxies selected from the VIKING-VIDEO-CFHTLS overlap region. Panel (a) shows all galaxies, Panel (b) shows remaining galaxies after morphological cuts have been imposed and Panel (c) shows the sample after both morphological cuts and an 8σ detection significance threshold in Y and J is applied. The red dashed line shows a conservative colour selection constraint, which rejects all remaining galaxies.

at query time. Figure 4.6 shows linear and log-linear histograms of VIKING-SDSS matched objects binned by angular separation.

Number counts are heavily peaked at angular separations of $\theta_r \lesssim 0.7$ arcsec with sparsely populated tales indicating that a small fraction of objects have been falsely matched between surveys. Objects with positional offsets of $\theta_r \leq 0.7$ arcsec are assumed to be associated and are rejected as quasar candidates. Some optically detected sources will remain in the catalogue and will be rejected via visual inspection, as discussed in later sections.

A further few per cent of sources can be rejected by placing limits on inter-band astrometric offsets. Objects with a redshift have no detectable proper motion and so candidates with positional offsets larger than 0.7 arcsec between any two bands are rejected.

Finally an automated routine flags objects which are located within regions around bright stars affected by ghosting and diffraction spikes (chapter 2). Flagged objects are visually inspected and those which are clearly affected by a bright neighbour are discarded. This completes the first pass over the candidate list, leaving 1.0×10^4 objects.

At this stage candidates that are too faint in the Z-band to be detected by the CASU source extractor undergo independent aperture photometry via the Image Reduction and Analysis Facility (`IRAF`) task `phot`. To ensure consistency with the CASU measured magnitudes, counts are measured in CASU's default 2 arcsec diameter aperture. All other parameters are taken, for the relevant detection, from the VSA. Substituting terms

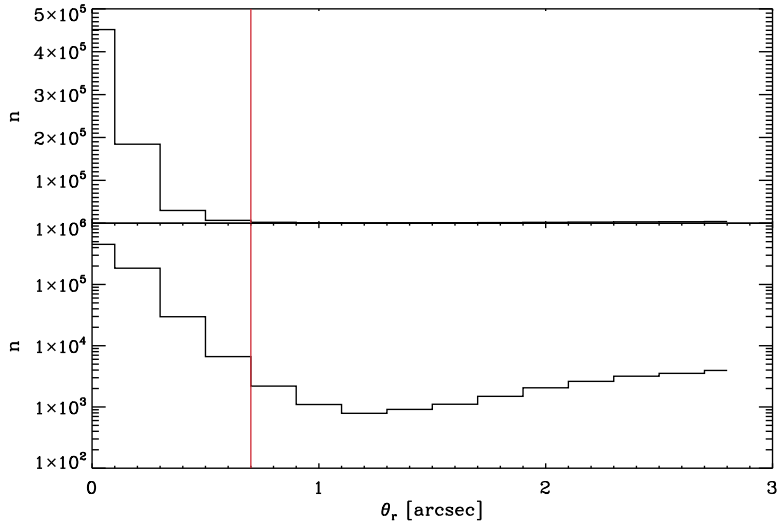


Figure 4.6: VIKING shares a region of overlap with the SDSS in the ngc. The figure shows the linear and log-linear histograms, binned by angular separation θ_r, in intervals of $\Delta_{\theta_r} = 0.1$ arcsec, for objects matched between the two surveys. Such objects can be rejected since $z \geq 6.5$ quasars are expected to be dark in optical passbands. The histograms are heavily peaked at separations of $\theta_r \leq 0.7$arcsec. A sparsely populated tail suggests that a small fraction of objects have been wrongly matched. Objects at separations shortward of the red line at $\theta_r \leq 0.7$arcsec are rejected.

into equation 3.1.3 gives magnitudes in the asinh form as,

$$m' = -1.08574\left[\operatorname{asinh}\frac{(f/f_0)}{2\mathrm{b}} + \ln \mathrm{b}\right], \tag{4.0.4}$$

where the (f/f_0) is the extinction corrected ratio of the observed count rate to the zero-point count rate. Taking into account the attenuation of flux via extinction in the Earth's atmosphere at airmass X, with extinction coefficient $\kappa(\lambda)$, and correcting for the loss of light, A, due to the finite size of the extraction aperture, (f/f_0) is given for the object photon count, C_{obj}, at the zero-point magnitude p, for the relevant exposure time t_{exp} by,

$$\frac{f}{f_0} = \frac{C_{\mathrm{obj}}}{t_{\mathrm{exp}}}\, 10^{-0.4\,[p-\kappa\,(\mathrm{X}-1)-A]}, \tag{4.0.5}$$

A more detailed discussion on the method of aperture photometry and of the variables needed to make such a measurement is given in chapter 5.

All the remaining candidates are shown in figure 4.1 (e). Even if it were possible to secure the necessary telescope time to follow-up these candidates, doing so would represent an inefficient use of scientific resources. At this stage the conventional approach is to heavily confine the the area of searchable multi-colour space and work backwards,

prioritising candidates based on some measure of their distance from the stellar and galactic loci, perhaps by using brightness and signal-to-noise properties as (usually qualitative) secondary constraints.

4.1 Colour selection

Take as an illustrative example the first follow-up of VIKING quasar candidates, which was scheduled over three nights in June 2011. As a rough estimate one might expect this to be sufficient to make $\sim$ 100 observations. This will seem better justified in the following chapter by referring to the ESO exposure time calculator (ETC). Prior to follow-up, there was expected to be $\sim 300\,\mathrm{deg}^2$ of data available from which to choose candidates. This allows a first broad-brush selection constraint to be defined e.g. that the candidate catalogue must be limited to ~ 1 object per $3\,\mathrm{deg}^2$. To place further, similar benchmark constraints it is necessary to look in more detail at contamination of quasar colour space. The word 'benchmark' is used here to emphasise that the following approach will be conservative i.e. that the colour constraints will be tighter than are perhaps necessary. In a real survey it is often useful to relax constraints e.g. so that one can consider errors on candidates around the selection cutoff boundaries.

Referring back to figure 4.5 (c), most galaxies are concentrated within the main locus. There are perhaps half a dozen sources lying sufficiently far from the galaxy locus to be mistaken for potential candidates. Assuming that this distribution is typical for all VIKING galaxies, then in the worst case scenario all outliers from the locus would be considered for followup. This suggests, perhaps ~ 10 extra-galactic contaminants per deg^2, although this is clearly a weak constraint. Better constraints can be placed on the extent of extra-galactic contamination when candidates are identified spectroscopically. For now, it is safer to work conservatively by placing a set of benchmark colour cuts such that all galaxies are rejected. These cuts are shown by the red dashed line in figure 4.5.

Figure 4.7 plots the simulated cool-star catalogue described in Chapter 3 in ZYJ colour-colour space. Stars are represented as black points and contours. Plotted in blue is a catalogue of ~ 200 simulated $6.5 \leq z \leq 7.5$ quasars. Each source has been 'detected' at the $8\,\sigma$ level in Y and J and was drawn from a larger catalogue of ~ 1000 quasars with a uniform distribution in redshift. Given the colour selection space constraints derived from the galactic sample above, the task here is simple; to refine these criteria to allow ~ 1 interloper per $3\,\mathrm{deg}^2$ while simultaneously ensuring that the vast majority of quasars lie within these constraints. One must also ensure that the revised constraints do not constitute a relaxation of the above constraints.

The revised selection region is described by equation 4.1.1 and shown in figure 4.7 by the dashed line, which encloses ~ 500 interloping stars. As required, this corresponds

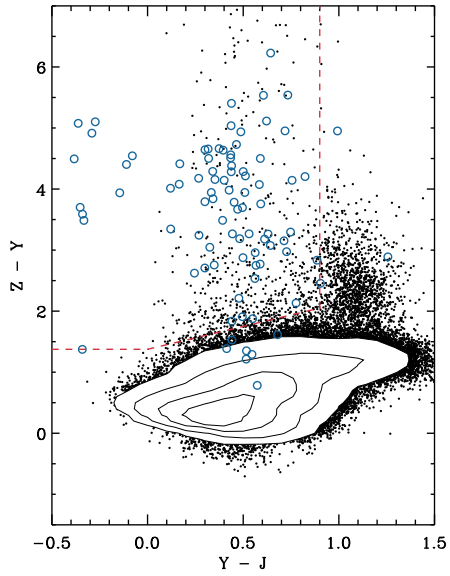


Figure 4.7: The VIKING Z, Y, J colour plane for simulated cool stars (contours and points) in the VIKING catalogue. Also shown are $\sim$ 200 simulated $6.5 \leq z \leq 7.5$ quasars (open circles) drawn from a uniform distribution in redshift and 'detected' at the 8σ level in Y and J. The red dashed line defines a benchmark colour selection region that encloses $\sim$ 1 object per 3 deg^2 as was required by the conditions of our follow-up campaign.

to roughly $\sim$ 1 interloper per 3 deg^2 over the entire 1500 deg^2 of the VIKING field.

$$\begin{aligned} Z - Y &\geq 1.375 \\ Y - J &\leq 0.900 \\ Z - Y &\geq 0.750\,(Y - J) + 1.375 \end{aligned} \tag{4.1.1}$$

By comparison with a $\sim$ 1 deg^2 patch of VIKING data, it was shown in chapter 3 that the modelled cool-star component of the VIKING catalogue, provides a realistic representation of the colour and magnitude distributions of the true populations. It is interesting to see how the above colour cuts, derived from the modelled catalogue, perform when applied to the $\sim$ 290 deg^2 of data available in June 2011. As well as applying these colour cuts the data-set was also subject to the detection significance, morphological and more general constraints applied in previous sections and each object was visually inspected in VIKING image frames to ensure the integrity of the detection in each passband (section 4.5). Figure 4.8 plots those objects falling within the benchmark colour constraints. In total there are 133 selected sources, which compares well with the $\sim$ 1 object per 3 deg^2 predicted by the modelled catalogue.

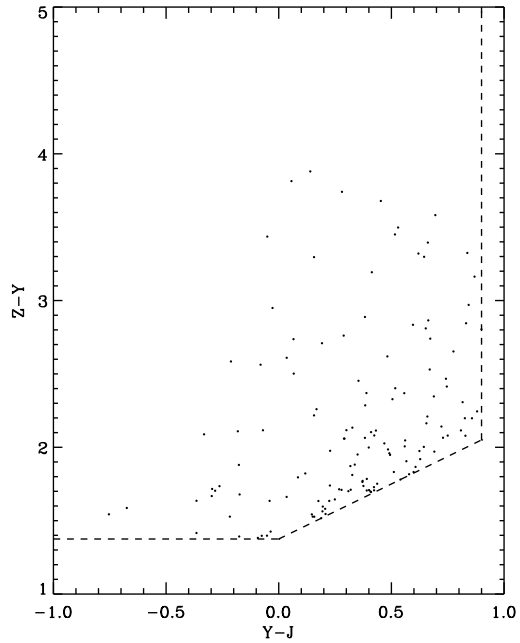


Figure 4.8: ZYJ colour plane for quasar candidates drawn from the $\sim 290\,\mathrm{deg}^2$ of data in the VSA as of June 2011. A series of constraints including, detection significance, quality control, morphological and colour, have been applied as described in the text. The constraints limit the candidate list to $\sim$ 1 object per 3 deg^2, as required by the follow-up strategy.

4.2 Probabilistic candidate ranking

Although colour selection has proven extremely successful in the past it will struggle to remain worth while into the next decade as very large area surveys ($\Omega \sim 20,000\,\mathrm{deg}^2$) with depths comparable to today's state of the art wide field surveys become common place [e.g. 69, 132, 133]. For rare object searches, the huge quantities of data necessitate a change of tack and several authors have already started to adapt to this change, by formulating candidate ranking schemes in a robust mathematical framework designed to make efficient use of computational resources [e.g. 127–129, 134].

As a simple example of such a ranking scheme, consider the Monte Carlo realisation of the colour space distributions of stars and quasars shown in figure 4.7. If one were to place a grid over the simulated colour space, then according to the simulation the probability of finding a real quasar in a given grid cell is $P_q = N_q/(N_q + N_s)$, where N_q and N_s are the number of simulated quasars and stars found in that grid cell. Assuming that the model parameters describe the true distributions exactly, then in averaging over increasing numbers of Monte Carlo realisations, the calculated probability would approach the true probability.

Thus one may rank a candidate on the probability of finding a simulated quasar in the region of colour space it occupies. This is an adequate starting point for a candidate

ranking scheme but it is not the finished product, since it neglects to consider all the information at ones disposal.

One obvious improvement would be to incorporate the extra information supplied by the measurement errors on each candidate. This random photometric error acts to perturb the true flux of a candidate to the flux that is actually observed.

For faint objects such as high-z quasars measurement errors are well known to be dominated by the uncertainty in the sky background estimation, which is typically Gaussian distributed. In other words the scatter induced in the true flux of a candidate by a measurement error is governed by a Gaussian probability distribution. This fact provides the opportunity to determine the probability that an object at a point in observed colour space has been scattered from some other point in true colour space, or more specifically, points occupied by the stellar and quasar populations.

This is best quantified by the likelihood, which is defined as the probability of obtaining a particular data set while assuming a certain hypotheses is true, or in the present context, the probability of a candidate having an observed colour assuming that it has originated from the stellar or quasar loci.

The likelihood shall be defined more formally below but first it is useful to note that unlike the Monte Carlo approach, which assigns a pre-determined probability to a region of colour space, the value of the likelihood is actually dependent on the data set itself.

Both approaches give a valid measure of probability and they may be combined to give a further measure via the use of Bayes' theorem. Initially the best constraint is given by ones prior belief, for instance the relative numbers of stars and quasars expected at given point in colour space. The crux of the Bayesian philosophy is in the use of evidence, or data, to weight the initial hypothesis as the evidence accumulates.

The first Bayesian algorithm designed specifically for high-z quasar candidate ranking was put forward by Mortlock et al. [130] and employs Bayesian model comparison techniques, in which candidates are ranked on the likelihood that they have scattered from the true (error-free) stellar and quasar loci. The technique has worked well for near-infrared searches in the UKIDSS-LAS and a similar approach is followed here.

Following Mortlock et. al. two distinct hypotheses are put forward for examination; that a given candidate is a high-z quasar or a cool-star, denoted Q and S respectively. As shown previously in this chapter, significant contamination from galaxies can be avoided via morphological and detection significance constraints. In this case, the extent of galactic contamination will be known post ranking and further measures to reduce such contamination can be investigated as and when the need arises.

In the following, we switch from working in colour space to work in flux space. The reason for this change is that the use of colours unnecessarily washes out informa-

tion on source brightness. This is particularly problematic in the case of high-z quasar selection since, as shown in figure 3.5, bright quasars are rarer than faint quasars. Furthermore the use of fluxes over magnitudes and colours allows for the correct interpretation of negative fluxes, which are simply a side effect of the more general problem of random measurement error, which has already been discussed.

Given a data set $\mathbf{D} = \{f_1, f_2, ..., f_{N_D}\}$ containing N_D flux measurements of a given candidate, the interpretation of probability as the confidence in the hypothesis Q is quantified by Bayes' theorem as

$$P(Q|\mathbf{D},\mathbf{I}) = \frac{P(\mathbf{D}|Q,\mathbf{I}) \times P(Q|\mathbf{I})}{P(\mathbf{D}|Q,\mathbf{I}) \times P(Q|\mathbf{I}) + P(\mathbf{D}|S,\mathbf{I}) \times P(S|\mathbf{I})}. \tag{4.2.1}$$

where all probabilities are conditional on the available background information contained in the set $\boldsymbol{I}$, that is to say, there is no absolute probability. In this case the background information is simply that the source under consideration has been detected at all.

To appreciate how equation 4.2.1 relates to the problem in hand a few definitions are useful; the subject of equation 4.2.1, which is hereafter termed P_q, is the posterior probability and represents the level of confidence in a hypothesis in light of the data. In this case it is the confidence that a particular candidate is a high-z quasar given its measured photometry. The terms $P(Q|\mathbf{I})$ and $P(S|\mathbf{I})$ are known as the priors, since they represent the state of knowledge *prior* to consideration of the data. Our anterior knowledge of the problem is provided by the Monte Carlo simulation which gives the surface density of objects from either class having the measured properties. The priors are then scaled by the likelihoods $P(\mathbf{D}|Q,\mathbf{I})$ and $P(\mathbf{D}|S,\mathbf{I})$, which describe the chance of obtaining the data given the hypotheses, or more specifically the probability of obtaining the measured photometry, from the intrinsic photometry described by Q or S and some underlying distribution of photometric scatter.

With a suitable model M, to describe the hypothesised distribution of intrinsic photometry, one can assume that the observed flux $f_{D,B}$ in the arbitrary bandpass B has been perturbed from its intrinsic flux $f_{M,B}$ as a result of photometric scatter $\sigma_{D,B}$, which is Gaussian distributed. The likelihoods may then be computed over N_B bands via an iterative process as follows

$$P\left(f_{D,B} \,|\, f_{M,B}\right) = \prod_{B=1}^{N_B} \frac{1}{(2\,\pi)^{\frac{1}{2}}\,\sigma_{D,B}} \exp\left\{ -\frac{1}{2}\left[\frac{f_{D,B} - f_{M,B}}{\sigma_{D,B}}\right]^2 \right\} \tag{4.2.2}$$

which is a Gaussian distribution with a diagonal covariance matrix as appropriate where inter-band noise correlations are negligible.

As in chapter 3 the quasar population model can be constrained from information at

lower redshift, where the photometric properties and number densities of quasars are well known. Mortlock et. al constrain the stellar population via a maximum likelihood fit to the real data. In the case of VIKING, this was not possible with the limited data in the initial release. An adequate substitute is the simulated stellar population derived in chapter 3, which was shown to predict the properties of the VIKING catalogue well.

There are however, some important differences between the catalogues modelled previously and those used for the probabilistic ranking. The following section describes these differences and tests the ranking algorithm on simulated test-set data to compare the approach to the traditional colour selection strategy.

4.3 Testing the ranking algorithm

In equation 4.2.1 the modelled catalogues are represented by Q and S, which reflect the intrinsic rather than the error perturbed distributions of each population. In chapter 3 catalogues were constructed to reflect the error perturbed distributions so for the purpose of probabilistic ranking each catalogue is re-constructed omitting Gaussian photometric errors. A second notable difference is that the new catalogues extend well beyond the VIKING detection limit. This ensures that faint interloper candidates, i.e. candidates which have found their way onto the candidate list as a result of photometric scatter, are dealt with fairly. The resulting catalogues then contain the photometric distributions, intrinsic to high-z quasars and cool-stars, which will be used to form the likelihoods.

The other ingredients are the prior probabilities, which describe the number densities of each population in flux space. This information is contained in the 3-dimensional density histograms of the catalogues described above. Thus, the resulting models of the cool-star and high-z quasar populations exist in two components, $\mathbf{Q} = \{Q_{\rm C}, Q_{\rm H}\}$ and $\mathbf{S} = \{S_{\rm C}, S_{\rm H}\}$, where subscripts C and H refer to the catalogue and histogram components respectively. Once these steps have been completed the algorithm proceeds in the following fashion

(a) A high-z quasar candidate with measured photometry $\mathbf{D} = \{f_{\rm Z}, f_{\rm Y}, f_{\rm J}\}$ is drawn from the list of candidates.

(b) $Q_{\rm H}$ and $S_{\rm H}$ are sampled at the relevant flux space bins. The occupancies of these bins are the values given to the prior probabilities $\mathrm{P}(Q_{\rm H}|\mathrm{I})$ and $\mathrm{P}(S_{\rm H}|\mathrm{I})$.

(c) The data $\mathbf{D}$ is compared to $Q_{\rm C}$ and $S_{\rm C}$ via the likelihoods. For a modelled catalogue of $N_{\rm C}$ sources, each with intrinsic flux $f_{\rm C,B}$ in a given band, an iterative procedure is used such that $\mathrm{P}\left(f_{\rm D} \mid f_{\rm C}, \mathrm{I}\right) = \Sigma_{\rm C=1}^{N_{\rm C}} \Pi_{\rm B=1}^{N_{\rm B}} \mathrm{P}\left[\left(f_{\rm D}\right)_{\rm B,C} \mid \left(f_{\rm C}\right)_{\rm B,C}, \mathrm{I}\right]$

(d) Bayes theorem is employed to assign a probability to the candidate.

(e) The process is repeated for all candidates and finally candidates are ranked on their probabilities.

In order to test the ranking algorithm, 50 of the input quasars were omitted from the construction of Q_C and Q_H and used instead in the construction of a further catalogue containing 10^4 simulated VIKING observations of quasars in the redshift range $6.5 \leq z \leq 7.5$ (i.e. including Gaussian photometric scatter). This catalogue was then used as a test set, whereby each simulated observation was run through the ranking algorithm and given a quasar probability based on its simulated photometry. The results of this exercise are shown in figure 4.9 where the test set is plotted in the VIKING ZYJ colour plane, each object is colour coded according to its probability. The black line shows the average evolution of the quasar locus as a function of redshift. The dependency of quasar probability on photometric scatter is clear, most quasars around the extremes of the distribution have large photometric scatter and consequently are almost always better explained as being stars.

In producing figure 4.9, the only pre-requisite is that quasars must have representable logarithmic Y- and J-band magnitudes. For typical quasars, probabilistic ranking becomes more reliable as the signal-to-noise increases. A further use of the test-set quasars is to investigate the relationship between signal-to-noise and P_q. Figure 4.10 plots the signal-to-noise of each test set 'detection' in Y and J, colour coded according to quasar probability. Most well detected objects are highly ranked, but there are reasonably well defined drop-offs in P_q evident along both axes towards lower signal-to-noise. This allows threshold detection limits to be defined, below which, quasar selection via probabilistic ranking will become unreliable. In this instance the thresholds are shown by the solid lines in figure 4.10 and are taken to lie at $S/N \leq 7\sigma$ in both Y and J.

One final point worth noting is that the ranking algorithm is necessarily biased towards the highest redshift quasars. This is inevitable since the higher the redshift the further the object lies from the stellar locus. At $z \simeq 6.5$ the Lyα transition contributes significant flux in the Z-band and with colours of $Z - Y \simeq 1.75$, $z \simeq 6.5$ quasars are only marginally redder than the coolest M-dwarfs, a fact which is authenticated by figure 3.2. Consequently most $z \simeq 6.5$, quasars receive a low P_q and are unlikely to be selected with ZYJ photometry using this method. This is evident from figure 4.9.

The discovery of a significant sample of $z \lesssim 6.5$ quasars in the UKIDSS-LAS made significant use of *i*-band data from the SDSS [37, 73–76]. At $z \simeq 6.5$, quasars lie ~ 3.5 magnitudes redward of the coolest M-dwarfs in $i - Y$. In the near future the VST is expected to provide complementary optical data with a ~ 2.5 magnitude gain in depth over the SDSS. Over the first five years of operation, the VST will conduct three imaging surveys in various optical passbands. The Kilo-degree Survey [KIDS; 78] will be twinned with VIKING and will provide complementary optical imaging over the SDSS u, g, r, i passbands, allowing high confidence probabilistic selection of $z \simeq 6.5$ quasars.

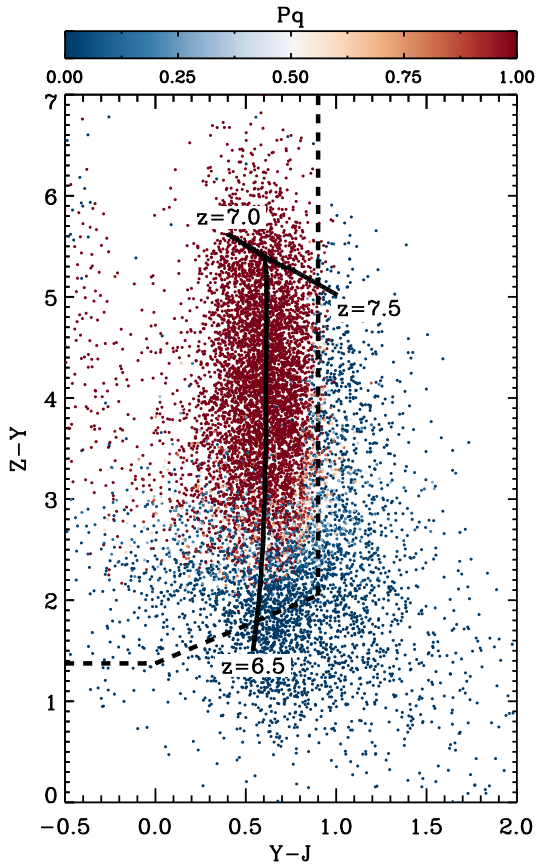


Figure 4.9: A simulated catalogue of $6.5 \leq z \leq 7.5$ quasars in the VIKING ZYJ colour plane. Each object is colour coded according to the probability P_q, that it is a quasar in the simulated redshift range. The thick solid line shows the average quasar redshift evolution track. Most objects lying significantly far from this track have large photometric errors and are consequently much better explained as being cool-stars. Quasars with $z \lesssim 6.6$ are only marginally redder than the coolest M-dwarfs and as such are mostly better explained as being cool-star interlopers. Forthcoming *i*-band data from the VST is expected to improve this situation.

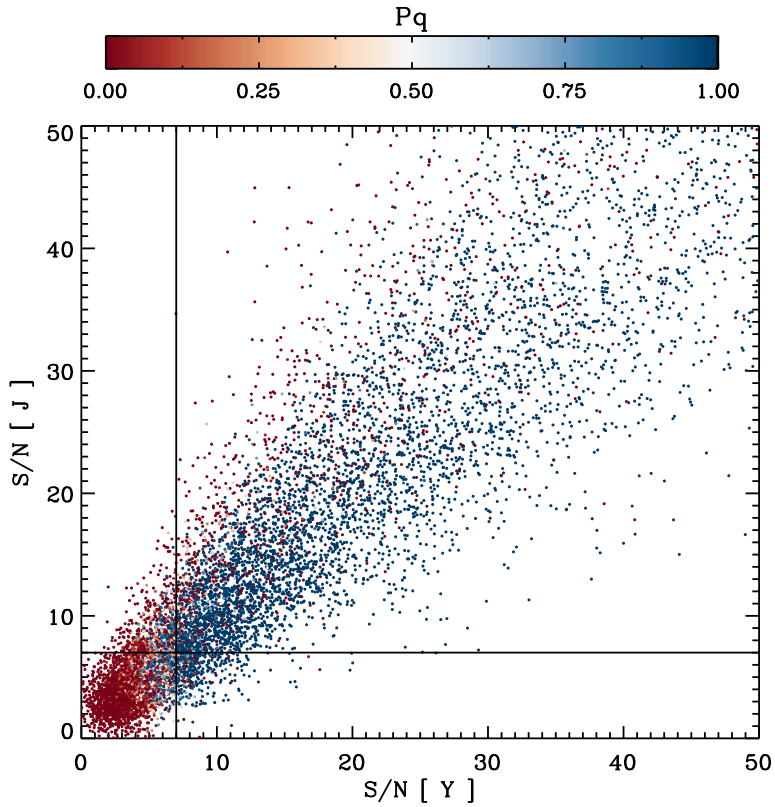


Figure 4.10: The YJ signal-to-noise distribution for a set of simulated quasars in the redshift interval $6.5 \leq z \leq 7.5$. Each object is colour coded according to its quasar probability P_q. Probabilistic ranking is effective down to reasonably well defined signal-to-noise thresholds shown by the vertical and horizontal lines at $(S/N)_Y = 8$ and $(S/N)_J = 8$.

In the most immediate future, the first VIKING quasar searches will be significantly biased against quasars with $z \lesssim 6.5$.

4.4 Completeness

In section 4.1 a set of benchmark selection constraints including colour, morphological and significance based discriminants, were defined to contain ~ 1 high-z quasar candidates per 3 deg^2. This corresponds to ~ 500 candidates over the entire 1500 deg^2 of the VIKING field that will require at least follow-up imaging, an achievable goal over the nominal five-year lifespan of the survey.

The estimated completeness contours for this selection strategy are shown as a function of J and redshift in figure 4.11 (a). The completeness function was determined by placing 10^4 simulated quasars (as described in chapter 3) into the VIKING selection space defined above. Completeness is defined here as the fraction of catalogued quasars (i.e. those detected by the CASU source extractor) of intrinsic J-band magnitude J and redshift z, with observables that pass the morphological, significance and colour constraints outlined above.

The total mean completeness of the survey is $\sim$ 60 per cent over the redshift interval $6.5 \leq z \leq 7.5$ and the magnitude interval $18.0 \leq J \leq 20.8$. The completeness is robust against moderate changes in the selection criteria that maintain the ~ 1 interloper per 3 deg^2 requirement.

The conservative selection constraints imposed on the simulated catalogue maintain a completeness that is comparable to but slightly lower than the highly successful SDSS and CFHTLS quasar searches in the optical and the UKIDSS search in the near-infrared [4, 125, 135, 136]. The difference can be accounted for firstly by our conservative selection constraints and secondly because completeness estimates in these searches do not consider the losses due to morphological rejection. With a mean completeness of $\sim$ 60 per cent, the quasar number count models derived in chapter 3, suggest that ~ 1.5 quasars in the interval $6.5 \leq z \leq 7.5$ will be selected from the first ~ 300 deg^2 of VIKING imaging, by the benchmark selection constraints.

Estimated completeness contours for the selection strategy based on probabilistic ranking are shown in figure 4.1.1 (b). Here the completeness is defined with the same morphological constraints as above, however significance constraints have been relaxed to accept candidates detected above the 7σ level in Y and J and colour constraints have been replaced by a threshold $P_q \geq 0.01$. This choice is somewhat arbitrary and part of the appeal of the probabilistic ranking method is that the choice of threshold P_q can be set according to the specific conditions of the search, available follow-up time being a good example. The total mean completeness is $\sim$ 67 per cent, representing a significant improvement over colour selection method. The difference is largely accounted

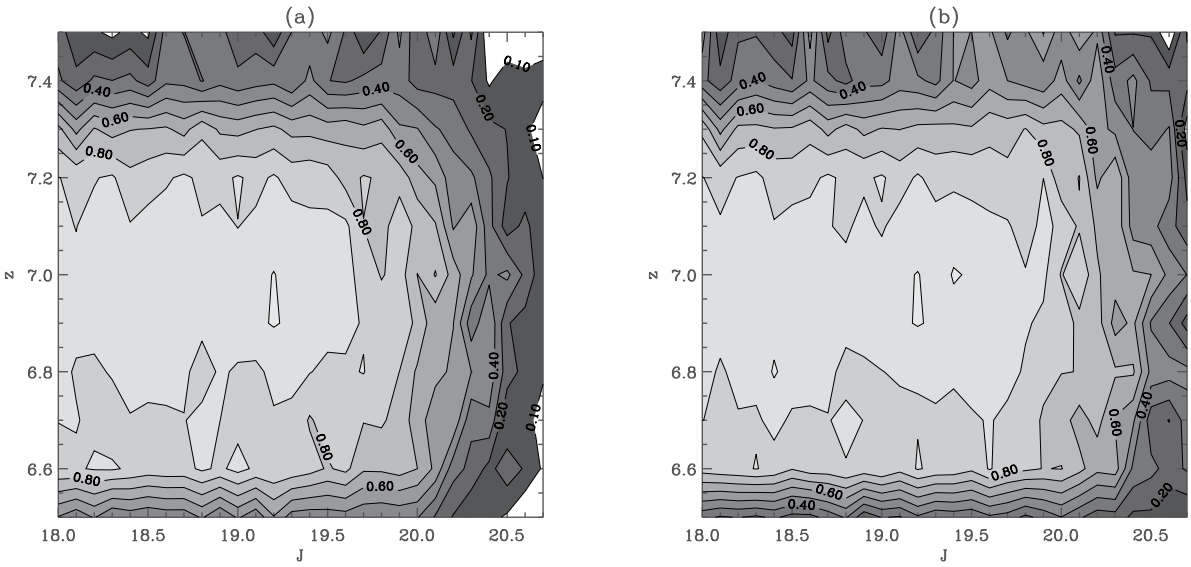


Figure 4.11: Panel (a): Estimated completeness for the benchmark colour selection constraints described by equation 4.1.1. Contours are plotted in J vs. z and increase in intervals on 0.1. The selection function accounts for loss of completeness from all significance, colour and morphological constraints described in the text. Panel (b): As for panel (a) but where colour constraints have been replaced for a benchmark $Pq \geq 0.01$.

for by the relaxation in detection significance thresholds. With a mean completeness of $\sim$ 67 per cent, the quasar number count models predict $\sim$ 1.7, $6.5 \leq z \leq 7.5$ quasars from the first $\sim 300\,\mathrm{deg}^2$ of imaging.

It is important to mention that the completeness will be further affected by a few additional factors not modelled here. Firstly, loss of completeness due to survey 'holes' caused by ghosts around bright stars; secondly, blending with foreground objects causing genuine quasars to move bluewards in $Z - Y$ and be rejected and finally, non-Gaussian errors and spurious detections causing increased contamination compared to the models.

The bright-star losses are expected to be relatively small, rejecting typically a few percent of the sky. Secondly, the losses from chance projection of foreground objects are also modest; at $Z \sim 23$ the surface density of objects is ~ 10 per arcmin^2, so if we assume (conservatively) real quasars will be rejected if closer than 2 arcsec to such an object, this will cause an inefficiency of 3.5 per cent.

Finally, there are many sources of non-Gaussian errors which may create spurious Z-dropouts, including moving objects such as asteroids, image persistence, hot pixels, etc. The VIKING observing strategy ensures that each object is observed at approximately 8 jitter positions on different pixels, and also that the J-band imaging is split into two epochs separated by weeks or months. Using this, it is expected that almost all spurious, time-variable and moving objects can be rejected by checking of individual data frames, though the precise number remains to be quantified in detail.

4.5 Selecting candidates

Candidate selection via Bayesian probabilistic ranking has already seen great success in the UKIDSS-LAS, where seven high-z quasars have been discovered in $\sim 2600\,\mathrm{deg}^2$ of imaging (see Section 4.4 for supporting references). Conceptually the approach taken by the UKIDSS team is no different to that taken here, however there are some notable differences regarding data input and implementation.

Firstly, UKIDSS searches have access to *i*-band data over most of the LAS field, the need for *i*-band data in the selection of $z \simeq 6.5$ quasars was discussed in section 4.3. Secondly the approach taken by the UKIDSS team in modelling the population of cool-stars was to fit directly to colour and magnitude distributions of the LAS imaging data, which differs markedly to the approach taken here. Although the models used here work well in characterising simulated test set data, they are yet to be tested in real world scenarios.

In practise then, selection was carried out independently by members of the VIKING working group and a combination of colour selection and probabilistic ranking was used to converge on the final candidate list. The work done during this thesis was concerned with probabilistic ranking. All candidates observed during follow-up were ranked probabilistically, regardless of the initial selection method. In the next chapter these candidates are discussed in detail, in the this section, selection via the ranking algorithm is discussed in more general terms.

At the beginning of this chapter a catalogue of 3.2×10^5 objects with quasar like colours was defined by a set of colour cuts laid down in chapter 3 and described by equation 3.3.1. By section 4.0.4 this catalogue was reduced by a factor 0.97 to 1.0×10^4 via a wide and varied series of cuts. Figure 4.12 (a) shows the remaining sources in the ZYJ colour plane after probabilistic ranking. Objects with $P_q \geq 0.01$ are depicted by open circles and colour coded according to their P_q. The power of the ranking algorithm is immediately clear. In total just 155 candidates are selected with $P_q \geq 0.01$.

Automated tasks have already been employed to remove either false detections (Chapter 2) or real objects which can be ruled out on the basis of e.g. optical detections, detectable proper motions, quality issues etc. At this stage there are possibly unknown issues still inherent in the data, as well as known issues that have been missed by the automated routines. To remove such detections each of the remaining candidates must undergo visual inspection. Figure 4.13 shows examples of a number of the most prominent issues found during this stage, details are listed correspondingly below;

(a) As discussed in chapter 2, bright star ghosting and diffraction spikes affect the photometry of objects lying in their immediate vicinity. Similarly bright stars produce local maxima around them which are interpreted as detections by the source extraction software. The size and position of halo ghosts around bright stars is

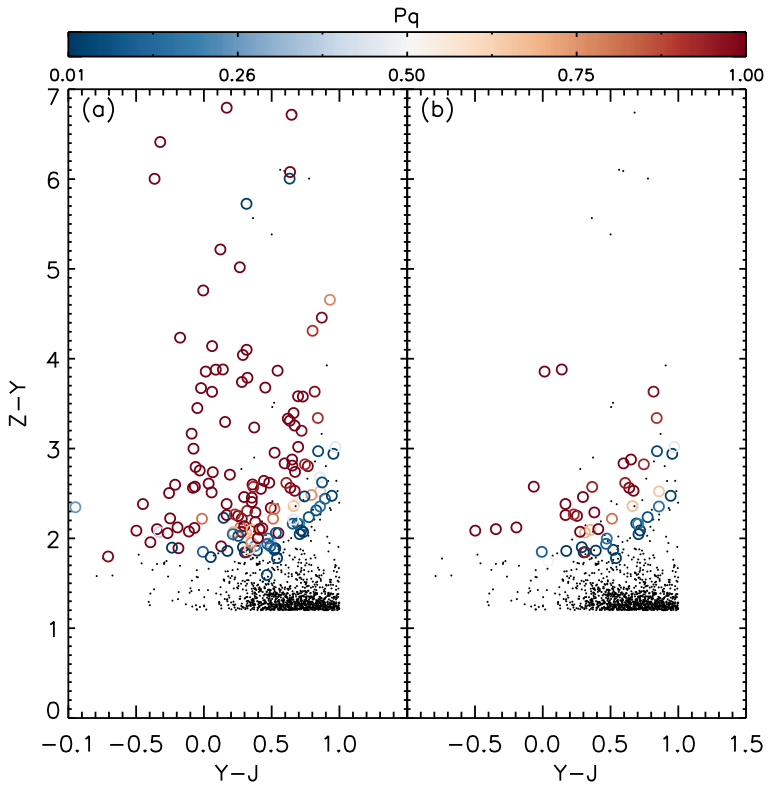


Figure 4.12: Panel (a): The candidate catalogue as of section 4.0.4. Each source has undergone probabilistic ranking and those with $P_q \geq 0.01$ are colour coded accordingly. Panel (b): As in panel (a) but after visual inspection of each source on VIKING image frames has led to the further rejection of spurious candidates.

predictable and most objects affected in this way were removed prior to ranking. The presence of diffraction spikes is a much more complicated problem and many diffraction spikes found their way through the automated routines designed to remove them.

(b) Large numbers of images were found to be affected by a similar pattern of noise in all consecutive images over different passbands. The reason for this inter-band correlation is unclear, but false multi-band detections on spatially correlated noise spikes can be clearly identified.

(c) Detector 4 is affected by a set of bad rows that do not correct well during image processing. False multi-band detections on spatially correlated bad pixels are not difficult to spot and can be easily removed.

(d) Close near neighbours may give incorrect photometry because light from multiple sources falls within the aperture. The true photometry of candidates affected by close neighbours cannot be easily reconciled and thus such sources are avoided.

(e) Detector 16 is affected by variable quantum efficiency on short timescales, which makes accurate flat fielding impossible. The top half of the detector bares the brunt of this unfortunate fault and the effect is more prominent at shorter wavelengths. In worse case scenarios science images show the mottled pattern seen in the figure, which can cause either false detections or false drop-outs. Detector 16 images are usually easily identifiable.

In section 4.0.4 some effort was made to remove candidates with SDSS counterparts found within 0.7arcsec of their VIKING positions. Any remaining optically detected objects are removed at this stage by visually inspecting candidate positions in SDSS cut-outs. Rejection of optical candidates and those affected by the issues outlined above, leaves the 50 objects shown in figure 4.12 (b). All remaining candidates are put forward for follow-up imaging.

4.6 Summary

The selection of high-z quasar candidates from the VSA via two contrasting methods has been investigated. A catalogue of objects with quasar-like colours was extracted from the VSA and a set of 'broad-brush' constraints applied to retain only the most promising objects. A region of $\sim 1\,\mathrm{deg}^2$ matched to deep optical and near-infrared observations from the VIDEO and CFHTLS surveys provided accurate optical-near-infrared colours which were employed to give robust star galaxy separation in this region. The sample was then investigated to place morphological constraints on VIKING point sources.

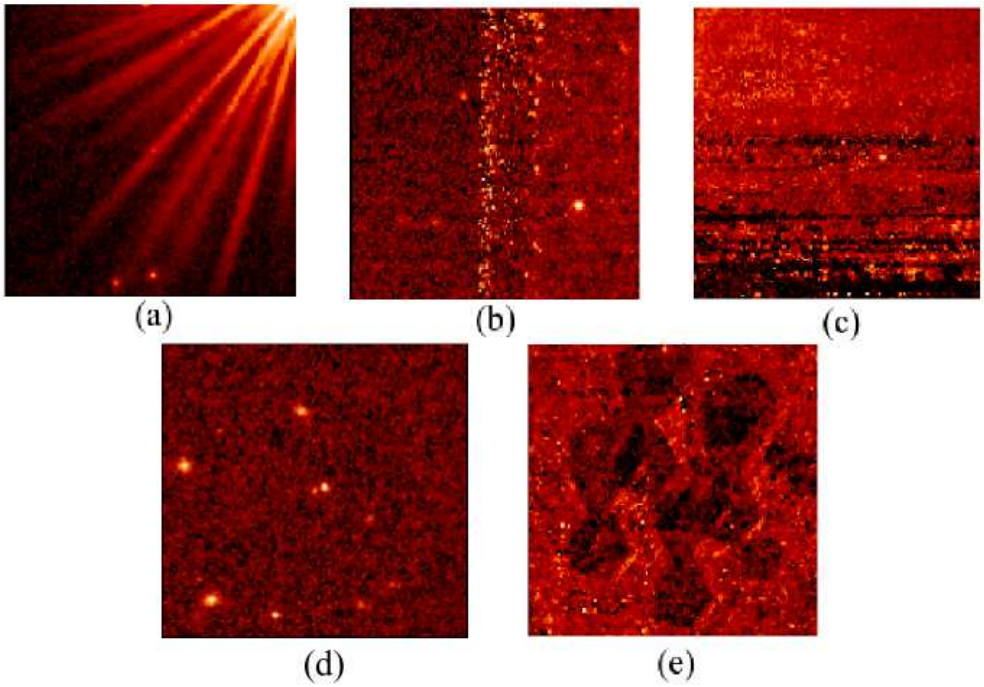


Figure 4.13: Visual inspection of highly ranked $(P_q \geq 0.01)$ quasar candidates reveal many spurious detections or drop-outs. Common issues are shown here and described in an itemised list, (a) - (d), in the main text. Candidates are positioned at the centre of each image.

A simulated catalogue of the cool star component of the VIKING survey was then employed to place some tight colour selection constraints on the quasar candidate list, limiting objects to a reasonable number for follow-up study.

For comparison, the mock VIKING catalogue was employed along with a mock quasar catalogue in a Bayesian model comparison strategy to rank objects on the probability that they were drawn from a population of quasars.

The completeness functions of the two strategies were then compared and it was found that probabilistic technique is the more complete. This strategy was employed to rank the quasar candidate list. Highly ranked candidates were put forward to the VIKING working group for follow-up observations.

CHAPTER 5

Photometric followup

Given any high-z quasar candidate, a low-resolution spectrum would be sufficient to confirm the true nature of the source without question. A low-quality spectrum of a typical candidate could be obtained in $\sim$ 30 minutes on an 8 m telescope. In the previous chapter 50 such candidates were earmarked for follow-up observations and since most of these are expected to be cool-stars scattered to quasar like colours, it is much more efficient at this stage to refine and extend our knowledge of the photometric properties of these sources in order to identify the false positives prior to spectroscopy.

The follow-up of VIKING quasar candidates was undertaken by a number of members of the VIKING working group. The 50 candidates defined in chapter 4 form a sub-sample of a larger set of interesting candidates put forward independently by members of the group, a fraction of which have since undergone follow-up imaging as discussed in the following chapter.

As a requirement of their selection, candidates are already reasonably well detected in both Y and J and thus most of the uncertainty in their photometry comes from the Z-band, in which they are necessarily faint. Thus follow-up in the Z-band is crucial for independent confirmation of the expected break shortward of the VISTA Y-band.

As was discussed in chapter 4, other than the Z-band, the *i*-band offers very useful information. At $z \geq 6.5$ quasars are all but dark in the *i*-band and as we shall see the expected significant contamination from M-dwarfs and to a lesser extent early L-dwarfs can be efficiently identified by a detection in a short *i*-band exposure. The combination of *i*- and Z-band observations thus have the potential to unequivocally rule out the Lyα break in the majority of contaminant objects in the sample, thus cutting the sample size down to such a level as to make spectroscopy a more realistic possibility. Having decided on a wavelength regime in which to observe the quasar candidates the next step is to decide on a suitable telescope and instrument.

5.1 The ESO New Technology Telescope

The ESO New Technology Telescope (NTT) Faint Object Spectrograph and Camera 2 [EFOSC2; 137] offers a z-band filter with a transmission profile very similar to the corresponding SDSS filter. The EFOSC2 i-band filter is slightly redder than the SDSS equivalent having a central wavelength and FWHM of 8024 Å and 1493 Å respectively. That is compared to the SDSS i-band which has a slightly bluer central wavelength of 7481 Å and a FWHM of 1526 Å. The EFOSC2 filter/detector combinations are henceforth referred to as $i_{\rm NTT}$ and $z_{\rm NTT}$.

EFOSC2 is permanently mounted at the Nasmyth B focus of the 3.58 m NTT at the La Silla observatory in Chile. The NTT is a facility that has been well used by members of the VIKING working group [73] and it is natural to investigate its potential to isolate $6.5 \leq \mathrm{z} \leq 7.5$ quasars.

As we have seen, the most common contaminants of quasar selection space are expected to be late M- and early L-dwarfs. Figure 5.1 plots the synthetic colours of quasars and cool-stars in various combinations of the $i_{\rm NTT}$, $z_{\rm NTT}$, Y and J bandpasses. The plots show that $\mathrm{z} \geq 6.5$ quasars are ~ 4 mags redder than the coolest L-dwarfs in $i_{\rm NTT} - \mathrm{Y}$, making this combination a powerful tool in establishing the high-z nature of any true quasar, or more likely, the Galactic nature most cool interlopers.

The coolest T-dwarfs can be significantly redder than the coolest L-dwarfs, with $i_{\rm NTT} - \mathrm{Y} \sim 6$. The quasar candidate list extends to magnitudes as faint as $\mathrm{Y} \simeq 20.5$, so for the coolest T-dwarfs one would expect a magnitude $i_{\rm NTT} = 26.5$ and like the targeted quasars, this renders the coolest T-dwarfs essentially dark in $i_{\rm NTT}$. Candidates have been selected with strong constraints in the near-infrared and most T-dwarfs will have been ruled out in the selection process by virtue of their red $\mathrm{Y} - \mathrm{J}$ colours in comparison to quasars. It would not be surprising to find that a small number of T-dwarfs still remain in the candidate list after follow-up imaging has ruled out the warmer stars. As has already been discussed, such objects are rare and a confirmed detection would constitute an useful scientific discovery in its own right.

Unsurprisingly, the $z_{\rm NTT}$, Y, J colour plane shown in figure 5.2, is very similar to the ZYJ colour plane. However an observation in $z_{\rm NTT}$, will provide important independent information on the presence or lack of a Lyα break blueward of the VISTA Y band. To confirm the break, an observation similar in depth to that of the original VISTA Z-band image is all that is required.

Candidate follow-up in $i_{\rm NTT}$ and $z_{\rm NTT}$ seems like a valid option in principle. To determine if such a campaign of follow-up is workable in practise, ESO provide an ECT [exposure time calculator, 138].

The late M- to early L-dwarf sequence extends to $i_{\rm NTT} - \mathrm{Y} \simeq 3.25$, thus a colour of $i_{\rm NTT} - \mathrm{Y} > 3.75$, i.e. undetected down to $i_{\rm NTT} \simeq 24.25$ at the 3σ level, would therefore

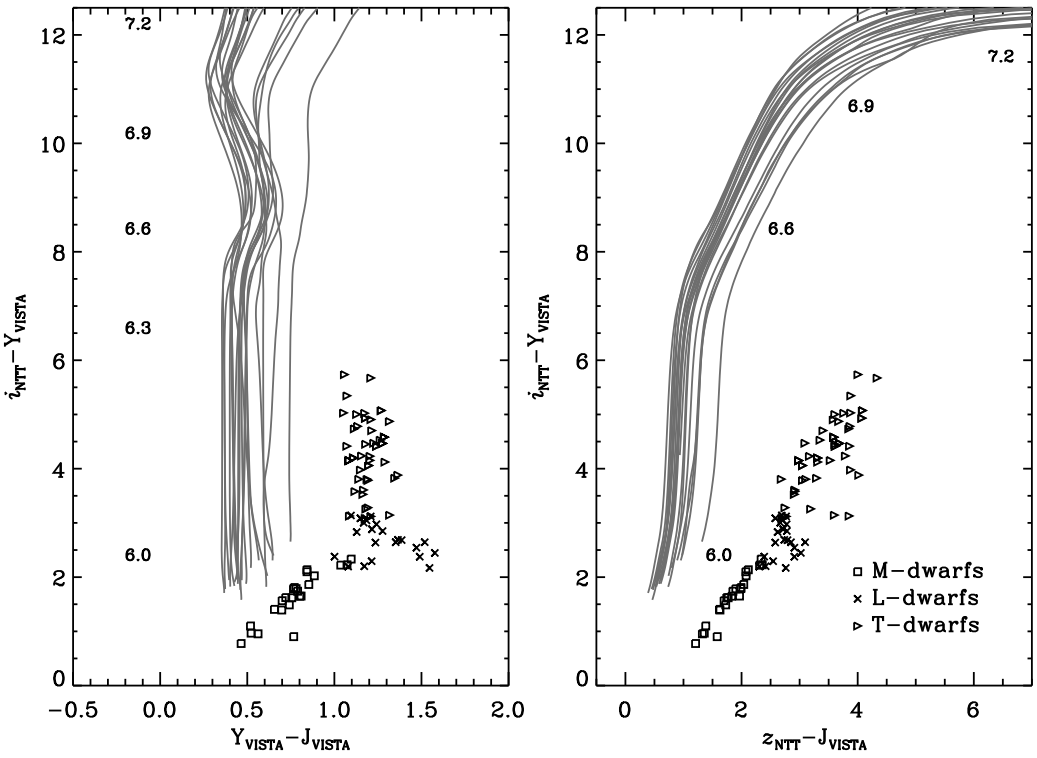


Figure 5.1: Left: Synthetic colours of cool stars and quasars in the $i_{\rm NTT}$Y, YJ colour plane. Right: The $i_{\rm NTT}$Y, $z_{\rm NTT}$J colour plane. z $\geq$ 6.5 quasars are significantly redder than M- and L-dwarfs in $i_{\rm NTT}$ − Y, which is good reason to obtain deep *i*-band imaging of high-z quasar candidates.

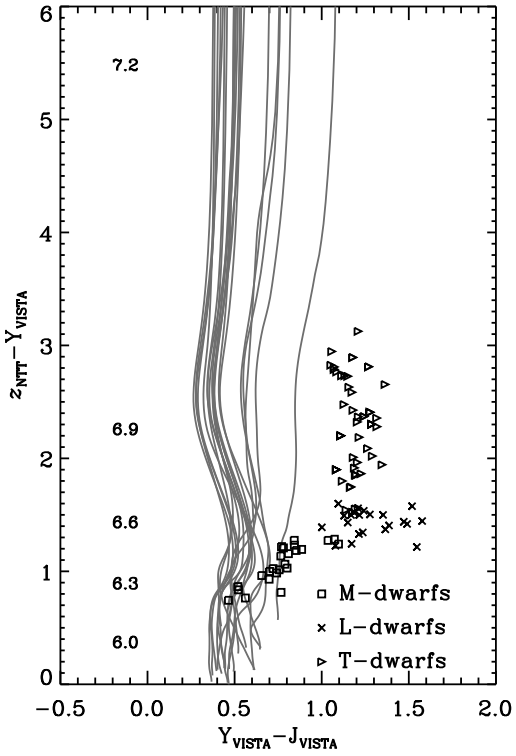


Figure 5.2: The $z_{\rm NTT}$ YJ colour plane. $z_{\rm NTT}$ imaging of quasar candidates will confirm the VIKING Z-band photometry.

signify an interesting object. Aiming to observe $\sim$ 100 of the most promising candidates put forward by the VIKING working group in $i_{\rm NTT}$, while optimistically anticipating 3 high-z quasars (work in chapter 3 suggest $\sim$ 1.5 to be a more realistic haul) requiring $z_{\rm NTT}$ observations, it is possible to determine the telescope time required from the ESO ETC.

Assuming observations are made at an airmass of 1.6, in seeing conditions of 0.8 arcsec, seven days from a new moon (i.e grey time), EFOSC2 can reach the required $i_{\rm NTT}$ limits in 625 seconds. Including overheads e.g. pointing, telescope preset, detector read-out, this corresponds to 15 minutes per target.

In similar conditions a 600 sec exposure in $z_{\rm NTT}$ will reach 24.0 mags at 3σ, which is all that is required to confirm or rule-out the Lyα break. Assuming three quasars are in the sample, each of which will require $z_{\rm NTT}$ observations, the total telescope time required is 25.75 hours, or 3 nights.

A proposal for time was put forward to ESO in September 2010. Time was awarded over three nights in June 2011 on program 087.A-0655(A).

5.2 Observations

Given the faint nature of our quasar candidates, optimising signal-to-noise becomes a crucial factor in confirming a detection or otherwise. Steps can be taken to increase signal-to-noise both at the imaging stage (as discussed here) and at the data reduction stage (section 5.3). Signal-to-noise is an increasing function of exposure time, but in a rare object search there is an obvious trade-off in maximising target coverage and ensuring that the most interesting candidates are well exposed. An observation strategy needs to be designed to balance these two factors.

5.2.1 Observation strategy

The noise inherent in photon counting comes from two distinct processes, the measurement of the signal itself, known as the source of object noise, and the noise from foreground light in the Earth's atmosphere, known somewhat ambiguously as the sky background. On a photometric night at La Silla, the I-band sky brightness is typically $\mathrm{I} \simeq 19.5\,\mathrm{arcsec}^{-2}$ [138], which in this case is several magnitudes brighter than our targets. Thus the object noise is negligible and the photon noise from the sky background dominates.

For a Poisson process the noise or uncertainty in the distribution is given by the standard deviation as $\sigma_{Poiss} = \sqrt{N}$, with N being the number of events or counts in a given interval. For astronomical observations the counts of interest are those associated with an astronomical source. For sky-limited observations the contribution to the total count from the source count, C_{obj}, is small compared with that from the sky C_{sky}, and the signal-to-noise attributed to the photon counting processes is $\mathrm{S/N} = \mathrm{C_{obj}}/\sqrt{\mathrm{C}_{\mathrm{sky}}}$. In considering the count rate rather than its integral, it is clear that the signal-to-noise increases with exposure time, $\mathrm{S/N} = \dot{C}_{\mathrm{obj}}\sqrt{t_{\mathrm{exp}}/\dot{C}_{\mathrm{sky}}}$.

Given the large number of candidates, it is desirable to devote longer integration time to the more interesting of them, which will typically be those that are undetected in i_{NTT}. The sum of multiple Poisson distributions is itself a Poisson distribution and thus the sum of multiple exposures on the same object retains the signal-to-noise characteristics of a single long exposure. The most effective use of our three nights then, will be to take short initial exposures on a large number of candidates in i_{NTT} on the first night and on subsequent nights, take longer i_{NTT} and z_{NTT} exposures on those that were undetected. Images can then be co-added to increase their effective exposure times. This allows for a much more effective observation strategy than would be offered by single long exposures on each candidate.

5.2.2 Imaging summary

The first VIKING quasar candidates were observed in $i_{\rm NTT}$ and $z_{\rm NTT}$ imaging by the EFOSC2 on ESO's NTT[1] over a three night period between 26th - 29th June 2011.

In total 129 science exposures were taken of 44 candidates[2]. All candidates were observed in $i_{\rm NTT}$ to depths of $\geq$ 23.5 mag at the 3σ noise level. At these depths z $\geq$ 6.5 quasars are dark in optical passbands so candidates showing a visible flux peak in these images were discarded, while those without detections were re-observed in $i_{\rm NTT}$ and $z_{\rm NTT}$.

Candidates are plotted in the VISTA ZYJ colour plane in figure 5.3. Error bars show the sky noise about the measured magnitudes with arrows representing lower limiting colours. Each candidate underwent probabilistic ranking prior to follow-up observation and each object is colour coded according to $\rm P_q$. With the final candidate list having been compiled from a number of independent methods (including probabilistic ranking; chapter 4) candidate rank was not used definitively at this stage to accept or reject candidates, but rather as guide to prioritising them.

In total just 6 candidates appeared undetected in $i_{\rm NTT}$ images that reached 3σ magnitude limits of $i_{\rm NTT} > 24$ (this has since been confirmed with aperture photometry). These objects are plotted as large open circles.

The $i_{\rm NTT}$ drop-out lying at $\rm Z-Y \simeq 1.0, Y-J \simeq 1.0$ is a faint (S/N $\simeq$ 6 in Y) low-priority candidate and would not normally have been observed but for a nightly period in which most high priority candidates were not visible. Given its low priority it was not observed in $z_{\rm NTT}$. Excluding this object there remain five strong $i_{\rm NTT}$ drop-out candidates which were subsequently followed up in $z_{\rm NTT}$ in order to confirm a break shortward of the VISTA Y-band. All were found to have a combination of colours in $i_{\rm NTT}$, $z_{\rm NTT}$, Y and J consistent with $6.5 \leq$ z ≤ 7.0 quasars.

The remainder of this chapter looks in detail at the reduction of the NTT imaging frames, subsequent measurements of the NTT photometry and presents the first results and tentative science from the VIKING high-z quasar search.

5.3 CCD Reduction

There are a number of factors which must be properly accounted for before accurate photometric measurements can be made from the NTT images. In particular the ob-

[1]Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 087.A-0655(A)

[2]Given the modest area covered by VIKING at this stage ($\sim$ 350 deg^2), it should be noted that the observation run was heavily dictated by the visibility of the patchy VIKING footprint throughout the night. Thus the observations presented here do not represent an exhaustive list of our highest priority candidates.

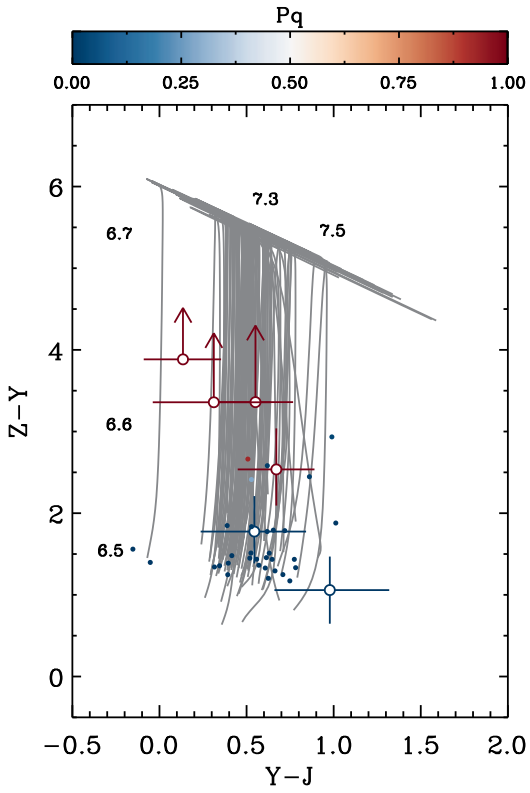


Figure 5.3: High-z quasar candidates in the VISTA ZYJ colour plane. Candidates are colour coded according to their probabilistic rank, P_q and over-plotted on a set of quasar colour evolution tracks labelled according to redshift. After NTT follow-up imaging, six candidates remain undetected in i_{NTT} and are plotted as large open circles. Error bars show the 1σ noise about the measured colours.

serving system is imperfect and leaves both noise and residual systematics in the images. These effects are collectively known as the instrumental signature and require careful removal.

5.3.1 Instrumental signatures

Instrumental signatures can vary from instrument to instrument but common to all is the thermal and electronic noise and the inhomogeneous illumination of the detector. These signatures are inherent over the entire CCD. Other more localised structures such as faulty pixels, rows or columns are of less importance here since they are unlikely to effect a single point source measurement over the 4 arcmin^2 field of EFOSC2. The same is true of satellite trails and cosmic ray hits, so the reduction requirements of the NTT images are the fundamentals described in this section.

Thermal noise

The laws of thermodynamics dictate that CCDs must have a non-zero temperature. Thus the high energy tail of the Fermi-Dirac distribution maintains a population of electrons in the CCD pixels without the need for 'activating' photons. This is known as the dark current and it can be isolated by taking an exposure with the camera shutter closed. The dark current can then be subtracted from science images. In practise, however cryostatic cooling of the detector in modern facilities such as the NTT renders the dark current insignificant (< 6 electrons pixel^{-1} hour^{-1} for the NTT) and it can be reasonably neglected. However with the NTT a dark frame is necessary since it is the only way to isolate the read-out signal bias.

Read-out signal bias

The bias is the electronic offset level due to the electronics and physical pixel characteristics. The bias is contained in an image taken with the shutter closed for the shortest possible exposure time. The bias frame can then be subtracted from the science frame. If a dark has already been subtracted there is not need to correct for bias since the dark already contains the bias signature.

Flat field correction

The final step in the basic reduction procedure is the flat field correction. The flat field corrects both for variations in pixel to pixel sensitivity and inhomogeneous illumination of the CCD. These signatures are inherent in the science images.

The inhomogeneity in the science images is isolated in an exposure of a homogeneously

illuminated area, which may be a flat screen, commonly referred to as a dome flat, or an image of the sky towards the zenith point just after sunset, as in a twilight flat. Flats also contain the bias and so this needs to be subtracted before the process of flat fielding can begin.

During twilight the sky brightness changes quickly, likewise dome lamps do not provide a constant flux. Thus flat field images must be normalised to unity and re-scaled to sky-level in the science images before application.

All of the above mentioned signatures contain noise, thus in correcting for instrumental signatures by subtracting the bias or dividing by the flat, the noise that is native to the calibration frames is copied to the science frames. In general calibration noise in modern instruments is low, but it is nevertheless important to retain control of it. Thus it is advisable to combine multiple calibration frames and average out the noise. The combined frames are known as the master-dark, -bias and -flat.

Image co-addition

Since we are on the subject of image combination, it is useful to describe the process that will be taken for both calibration and science frames in order to increase effective exposure time.

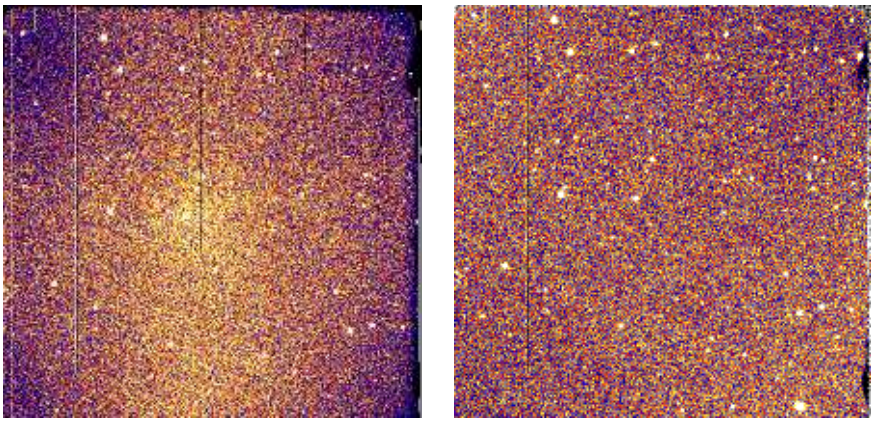
The two common options considered here are the median and mean combination. If a small number of images are to be combined then one should consider the median, since it is more resistant to outliers. However one should also be aware that for Gaussian distributed noise (i.e. for a high count rate), the noise level in the median is a factor ~ 1.25 larger than in the mean [139]. Thus in averaging science, bias and flat field images, the mean average is taken here.

To avoid the effects that outlying pixels can have in biasing the mean, a clipping algorithm is employed prior to combination. Mean pixel values are calculated excluding the extrema and deviations of greater than 2.3σ are rejected. The mean is again calculated from all remaining pixels and the clipping procedure is invoked once more. This process continues until all remaining pixels are within 2.3σ of the mean.

All stacked science and calibration frames are produced from mean combinations after undergoing this clipping algorithm. The entire process is undertaken by the `IRAF` task `imcombine`.

Reduced images

Once the master calibration images have been produced the reduction process is given as



(a) (b)

Figure 5.4: i_{NTT}-band short 5 sec exposure before (a) and after (b) bias correction and flat-fielding. Most obvious is the removal of the large scale gradient across the image.

$$\text{Reduced science image} = \frac{\text{raw science image} - \text{master bias}}{\text{master flat} - \text{master bias}} \tag{5.3.1}$$

Figure 5.4 shows an image of a bright 13th magnitude star taken with the NTT before and after reduction. Panel (a) shows the initial raw image in which subtle imperfections and large scale gradients across the image are visible. Panel (b) shows the reduced image in which the improvements are clear.

In this case, given the brightness of the target, a short 5 second exposure is all that was required to obtain a high signal-to-noise measurement. High-z quasars candidates are faint and require much longer integration times to achieve adequate signal-to-noise. Figure 5.5 (a) shows an image taken with a considerably longer integration time. The wavelike features in the image are known as fringing, which is inherent in all the images and needs to be corrected for in the images of quasar candidates.

5.3.2 Fringing

Fringing is an idiosyncratic variation in the background that is not a feature of the sky, but rather a thin-film interference pattern intrinsic to the CCD. The pixels in a CCD lie beneath a layer of intrinsic (i.e. undoped) silicon, which photons interact with to produce photo-electrons. The optical depth to the incident photons is a function of wavelength, red wavelengths having a lower interaction cross-section than blue wavelengths. When the optical depth is greater than twice the depth of the silicon layer, a fraction of photons are reflected from the pixel gates and interfere with oncoming photons. This interference pattern is what is referred to as fringing.

The contribution to the signal from fringing is additive and the pattern is usually sufficiently different from the stellar point source function (PSF) that it does not alter the stellar photometry significantly. The effect is instead to increase the background mak-

ing detections of faint objects difficult. Thus for faint sources such as high-z quasars, it is important to remove this artifact.

The night sky is dominated by telluric emission and this is reflected in the fringe pattern which is dominated by many monochromatic emission lines from the night sky. The chemical composition and temporal conditions of the atmosphere are, broadly speaking, stable. Thus the fringing pattern will remain approximately constant from image to image or indeed from night to night over a short observation run. This is also true of the fringing amplitude but only when conditions are photometric.

The fringe pattern may be extracted from a superflat. A superflat is an additional flat field image which is produced from a combination of the flat fielded science images. The superflat should be free of any real sky structures and as such, component images should not contain any large bright stars, galaxies or nebulosity and should be sufficiently dithered such that real sources are diminished via averaging during combination. The procedure is to ensure a uniform sky brightness by first normalising the science images about the mean, rejecting outliers. The normalised images are then mean combined with outlier rejection. The resulting superflat thus contains the fringing pattern. The superflat is then smoothed with a large Gaussian kernel (typically 200-300 pixels) giving the average sky background plus some inter pixel variation. The smoothed fringe frame is then subtracted from the original to return the normalised fringe pattern.

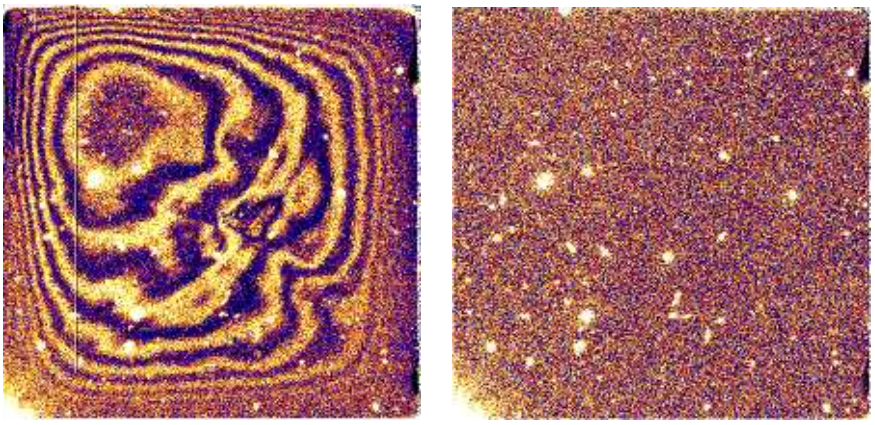
A fringe corrected science image may be obtained by subtracting the flat-fielded fringe pattern, scaled to the median sky background in the science image $\sigma_{\rm sky}$, from the fringe affected science image,

$$\begin{aligned}&\text{Fringe corrected science frame} = \\ &\text{fringe affected science frame} - \frac{\text{fringe pattern} \times \sigma_{\rm sky}}{\text{super flat}}.\end{aligned}$$

A fringe corrected science frame is shown in figure 5.5 (b).

5.4 Aperture photometry

With the images in science ready form, attention can now switch to photometric measurements. The concept of aperture photometry is satisfyingly simple. A star is enclosed within a digital aperture containing N pixels and pixel counts C are integrated to give a total flux. The average count of the background sky $C_{\rm sky}$ is also measured in a region free from any star light and the difference between the two gives the flux from the star. The asinh magnitude is then calculated via the following recipe,



(a) (b)

Figure 5.5: i_{NTT}-band long 600 sec exposure before (a) and after (b) fringe correction.

$$m_{inst} = -1.08574\left\{\operatorname{asinh}\left[\frac{(\Sigma_{\mathrm{i}=1}^{N} C_{\mathrm{i}}) - NC_{\mathrm{sky}}}{2\,\mathrm{b}\,f_0}\right] + \ln(\mathrm{b})\right\}. \tag{5.4.1}$$

In practice there are a number of more subtle steps that need to be taken before accurate photometry can be generated. The first of these steps places photometry on a meaningful relative scale.

5.4.1 Instrumental to standard magnitudes

In performing the fundamentals of aperture photometry described above on the NTT quasar candidate images, one will arrive at the instrumental magnitudes m_{inst}. The term reflects the fact that they are characteristic of the particular properties of the telescope, optical system and detector and further that they are subject to the effects of the atmosphere, which scatters and absorbs light on its way to the detector.

Instrumental magnitudes are therefore not directly comparable to measurements by other instruments or even to measurements by the same instrument at different times and are useful in only a few special cases. Thus the instrumental photometry must be placed on a standard system. This concept was met briefly in chapter 3, where synthetic photometry in the VISTA passbands was placed on the Vega system by defining the spectrophotometric standard star Vega as a zeroth magnitude star in all VISTA passbands. Here the premise is the same, i.e. all NTT magnitudes remain on the NTT system but are offset such that Vega is a zero magnitude star in all relevant NTT passbands.

By anchoring the photometry in a given passband to a common zero-point, the implication is that all instrumental magnitudes in this passband are offset from the standard by a constant magnitude. The task is then to find this magnitude.

If one were working at the top of the Earth's atmosphere then the zero-point p, would

be the difference between the magnitude of an object on the standard system and the same object on the instrumental system i.e. $p = m - m_{inst}$. The situation is made slightly more complicated by the atmosphere which scatters and absorbs light on its way to the detector; a phenomenon known as atmospheric extinction. The amount of extinction is a function of wavelength and path length. The path length is parametrised as airmass X, which is the ratio of column density along the line of sight to that at the zenith. Airmass is approximated as $X \simeq \sec\theta_Z$, where θ_Z is the zenith distance or the angular distance from the zenith to the line of sight. As the zenith angle of a particular star increases so too does the extinction, with the effect of dimming m_{inst}. The notion of the instrumental zero-point will therefore vary over the observation run.

The extinction coefficient κ gives the difference in magnitudes between a measurement taken at the top of the Earth's atmosphere towards the zenith and the same measurement taken at sea level, where $X = 1$. Away from the zenith, m_{inst} is dimmed by an amount $\kappa\,(X-1)$. In accounting for this dimming m and m_{inst} are related by,

$$m = m_{inst} + p + \kappa\,(X-1), \tag{5.4.2}$$

where p is the difference in absolute magnitude and instrumental magnitude at unit airmass. It is useful to define $p_{\rm x}$ to differentiate between p and the zero-point of an observation taken at arbitrary airmass $X = {\rm x}$, such that $p_{\rm x} = p + \kappa(X-1)$.

Given an m_{inst} measured at a known elevation, then X is a known quantity, and we are left with the three unknowns m, p and κ.

Conceptually this is resolved by taking intermittent measurements of bright standard stars for which the magnitude on the standard system is known. If an extensive set of standard stars have been observed over each night then m is also known and p and κ can be solved for via a least squares fitting, potentially resulting in photometry with sub per cent accuracy.

With the follow-up of rare objects there is a trade-off, one of course wants to observe as many candidates as possible and thus returning the telescope to a standard star field after every few candidate observations quickly becomes a hindrance. The observation strategy (section 5.2.1) was designed to some extent to make life simple, a candidate becomes interesting when it is undetected in $i_{\rm NTT}$ and this does not require extremely accurate photometry. What's more is that since quasar candidates are expected to be faint in $i_{\rm NTT}$ and $z_{\rm NTT}$ (typically $< 3\,\sigma$ above the noise), the sky background dictates that the uncertainty in the photometry will be $\gtrsim 30$ per cent, so an extensive set of standard star observations is really overkill for this particular application.

ESO provides a large data base of spectrophotometrically derived standard star spectra [140]. The Vega magnitudes of these stars in the NTT passbands [138] can be determined via synthetic photometry with equation 3.1.1 in a similar manner to that de-

Star name	i_{NTT}	z_{NTT}	Ra	dec
LTT-1020	10.8010	10.7360	01h 54m 49.68s	-27d 28′ 29.7″
LTT-4816	13.7584	13.8765	12h 38m 50.94s	-49d 47′ 58.8″
LTT-6248	11.1270	11.0538	15h 39m 00.02s	-28d 35′ 33.1″
LTT-9239	11.3062	11.2323	22h 52m 40.88s	-20d 35′ 26.3″

Table 5.1: Synthetically derived standard star NTT magnitudes on the Vega system.

scribed in chapter 3. Four such standards were chosen to calibrate the magnitudes of the quasar candidate observations. Each standard is given along with its synthetically derived i_{NTT} and z_{NTT} magnitudes and equatorial coordinates in table 5.1.

Observations of the four standard stars were made at the beginning, end and throughout each night. The resulting photometry was measured in large apertures containing ~ 100 percent of the source flux (see section 5.4.2 for discussion) and instrumental magnitudes were determined in the `IRAF` package `phot` via equation 5.4.1.

Figure 5.6 plots the instrumental zero-point, $p_{\mathrm{x}} = m - m_{inst}$, versus airmass for the i_{NTT} standard star observations on each night. Ideally equation 5.4.2 specifies, that each plot should be a straight line through p and equation 5.4.2 can be fitted for p and κ. However, for the small number of standard star observations made here, the results would be error prone.

Instead, equation 5.4.2 is fit with κ fixed to typical values. Atmospheric extinction can spike in volcanically active regions such as Chile where eruptions can increase the atmospheric aerosol content over periods of months [141]. Thus, without a robust measurement of κ during the observation run the best one can do is to look for long term monitored values published in the literature.

ESO themselves advise a value of $\kappa\,(i_{\mathrm{NTT}}) = 0.05$, which was measured on photometric nights during EFOSC2 recommissioning in April 2008 [142]. This value is broadly consistent with those measured by Tüg [143] at La Silla and more recently by Patat et al. [141] at ESO Paranal in both the i_{NTT} and z_{NTT} wavelength regimes. Thus equation 5.4.2 is fitted here with constant values of $\kappa\,(i_{\mathrm{NTT}}) = 0.05$ and $\kappa\,(z_{\mathrm{NTT}}) = 0.05$ leaving p as the only free parameter.

The straight line fits to the i_{NTT} zero-points are shown in figure 5.6 on consecutive nights. The fit residuals are shown as grey error bars in each plot. The zero-point at unit airmass, p_1, can be read directly from these fits and substituted into equation 5.4.2 to obtain calibrated magnitudes from the instrumental measurements. The results are summarised in table 5.2, which gives the value p_1 and associated residuals derived on each night in i_{NTT} and z_{NTT}.

The ~ 10 per cent uncertainty introduced by the sky background in the photometry of the quasar candidates would suggest that the estimated residuals in the zero-point fitting are insignificant in the calculation of candidate magnitudes. The exception to this

	night 1	night 2	night 3	night 3 (revised†)
i_{NTT}	25.08 ± 0.01	25.07 ± 0.02	25.02 ± 0.08	25.08 ± 0.00
z_{NTT}	24.31 ± 0.00	24.27 ± 0.02	24.28 ± 0.01	-

Table 5.2: Zero-points required to place NTT instrumental magnitudes onto the Vega system. Zero-points are derived via straight line fits to offsets between instrumental and absolute magnitudes of standard stars vs. airmass. Errors are fit residuals. †Revision of the i_{NTT} zero-point on night 3 was found to be necessary in section 5.4.4, see text for details.

is the zero-point determination in the i_{NTT}-band on night 3, where the residual translates into to a flux error of 7.5 per cent. Two outlying points would seem to be the cause of the increase scatter here. These points are associated with successive measurements of the same source within an interval of few minutes. The brightening of the zero-point translates an intermittent dimming of the source as measured on the instrumental system, the most common cause of which is passing cirrus. It would not be unreasonable to remove these outlying points in order to improve the fit. The possible systematics incurred by this issue are further discussed in section 5.4.4 and a decision as to whether a re-calculation of the affected zero-point is needed is deferred until then.

As directed by the observation strategy, objects that remained undetected in i_{NTT} in single imaging frames, were returned to on subsequent nights and further time was spent integrating in both i_{NTT} and z_{NTT} to build up large effective exposure times. It is desirable then to stack these images in order to produce a series of composites with improved signal-to-noise, but since each composite will incorporate images taken at differing air-masses on different nights, there is a need to account for these differences by placing each component image at an equivalent zero point.

The procedure is to multiply all component images relevant to a particular stack by a factor $10^{0.4\,(p_{\mathrm{x}}-p_1)}$, in this case p_{x} is the zero-point of the relevant component image at the airmass at which it was taken and p_1 is the zero-point at $X = 1$ on the chosen reference night.

5.4.2 Aperture radii

A second consideration is the size of the aperture. In space the image of a point source is diffraction limited and is therefore described by an Airy profile. For ground based observations the image is limited by the atmospheric seeing, which broadens the PSF. The seeing limited PSF is often well approximated by either a Gaussian, a modified Lorentzian or a Moffatt profile [144].

An aperture a factor 3 times the full width at half maximum (FWHM) of the PSF may seem advisable since it will typically contain $\sim$ 100 percent of the light [145]. For all but the brightest sources the wings of the stellar profile quickly become noise dominated

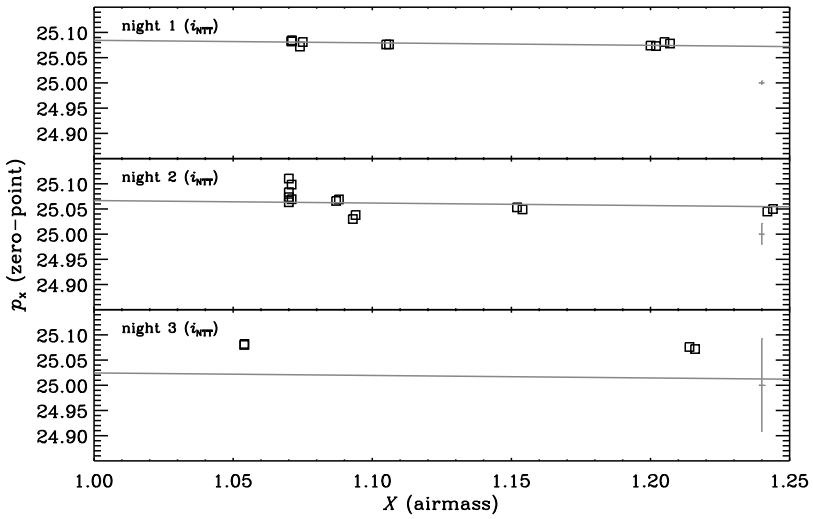


Figure 5.6: The variation in the $i_{\rm NTT}$ instrumental zero-point as a function of airmass on consecutive nights during quasar candidate follow-up. Zero-points are measured as the difference between absolute and instrumental magnitudes of bright NTT standard stars. The straight line fits assume an atmospheric extinction coefficient (slope) $\kappa = 0.05$ leaving the zero-point at airmass zero as the single free parameter. The fitted zero-point at airmass, $X = 1$ is substituted into equation 5.4.2 to place instrumental magnitudes onto the NTT Vega system. Similar fits were made for $z_{\rm NTT}$ standards.

and thus the optimum signal-to-noise is achieved by enclosing less than 100 per cent of the flux. An aperture that contains less than 100 per cent of the source flux will not give an accurate measurement of the source brightness and thus a correction term is added to account for the 'lost' flux (section 5.4.3).

As well as being related to the seeing the optimum aperture size is also a weak function of the source brightness becoming smaller for fainter objects [146, and references therein]. Prior to photometric measurement the brightness of a source is of course unknown, but for most well behaved PSFs the optimum is close the FWHM and is relatively stable in photometric conditions [146]. During the NTT observations the seeing conditions are written to the header units of the image files. The optimum aperture size can be related to the seeing assuming a Gaussian PSF with a standard deviation in pixels via,

$$\sigma_{\rm PSF} = \frac{\text{seeing (arcsec)}}{\text{pixel scale (arcsec/pixel)} \times 2.355}. \tag{5.4.3}$$

Patel [125] finds the optimum radius to lie at $1.7\,\sigma_{\rm PSF}$. During the NTT run the median seeing was ~ 0.9 arcsec, substituting this in to equation 5.4.3 with the NTT pixel scale of 0.157 arcsec/pixel gives an optimum aperture size of $1.7\,\sigma_{\rm PSF} \simeq 5.0$ pixels, which is the aperture size used for the following photometric measurements.

The average sky count per pixel $C_{\rm sky}$ is calculated within a halo surrounding the object

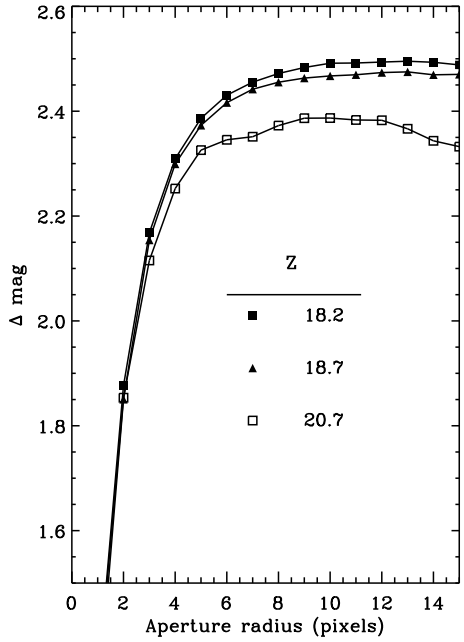


Figure 5.7: Growth curves: The magnitude differences within successively inclusive apertures for three stars on a single NTT imaging frame. The two bright stars follow the same canonical curve. The faint star follows the same curve until the sky-background dominates in the wings of the PSF. A correction factor can be calculated from the curves of the bright stars and applied to obtain the true magnitude of a star measured within a small aperture.

aperture. In this case the `centroid` algorithm supplied by `phot` is used to compute the average sky background as recommended by the package user guide [147], other authors have noted equivalent algorithms to give robust results [146]. One must ensure that the inner radius of the sky aperture is sufficiently distant from the object so as not to contain a "significant" fraction of the object count.

To determine exactly what is meant by a significant fraction of the object count, figure 5.7 plots the magnitude difference between successively inclusive apertures for three different stars from the same NTT imaging frame, otherwise known as the curves of growth. The two brightest stars, whose curves follow the filled squares and triangles, are the important ones here (the significance of the fainter star's curve will become clearer in section 5.4.3). If one approximates the PSF with a Gaussian function, then 99.7 per cent of flux lies within $3\sigma_{\mathrm{PSF}}$ of the flux peak. As the apertures approach the tails, the PSF is approximately constant over small distances and the rate of change in magnitude between successive apertures tends to zero, at which point the contribution of light from the star in remaining apertures is negligible. This occurs at a radius of ~ 15 pixels in the figure, which corresponds $\sim 3 \times$ FWHM of the PSF.

The PSF will vary with seeing so to err on the side of caution quasar candidate photometry is conducted with a sky-background level calculated in a halo of inner radius 20.0 pixels and width 10.0 pixels.

Aperture photometry was carried out on candidate followup images in the IRAF package phot via equation 5.4.1. The final step towards calibrated photometry is the aperture correction.

5.4.3 The aperture correction

Discussion in the previous section explained that star light falling a distance greater than $\sim 3 \times$ FWHM from the peak of the stellar PSF makes a negligible contribution to the total flux from the source. This was illustrated in figure 5.7 where the curves of growth of two bright stars originating from the same NTT imaging frame, were shown to follow the same canonical form tending to zero growth at ~ 15 pixels.

For fainter sources curves of growth deviate significantly from those of bright stars because the sky-background begins to dominate over the source flux in increasingly larger apertures. This is seen in figure 5.7 by comparing the growth curve of the fainter star to those of the bright stars.

To avoid such a state of affairs in measuring quasar candidates, photometry was performed in apertures of radius 5 pixels, where the growth curves of faint stars follow the expected curve. The obvious problem with such an approach is that there is significant flux at radii beyond 5 pixels from the PSF peak. It is therefore necessary to correct for the loss in flux inherent in using the 5 pixel aperture.

The underlying assumption of the aperture correction is that the PSFs of faint stars are exact replicas of those of bright stars. It follows that, in the absence of sky-background, their growth curves should possess equivalent functional forms. For a bright star, the ratio of flux in a large aperture containing ~ 100 percent of the bright star flux, to the flux for the same star in a small aperture gives the aperture correction factor in flux units, which is equivalently the magnitude difference between the two apertures.

Aperture corrections to the quasar candidates were calculated from bright stars in the corresponding frames by measuring magnitude differences between apertures of $3 \times$ FWHM and 5 pixels. Several frames lacked bright stars and in these cases aperture corrections were computed from the standard stars on corresponding nights.

The aperture corrections represent the final step in the computation of quasar candidate magnitudes. The results will be discussed with reference to the quasar search shortly. First it is interesting to see how the calibration of these magnitudes compare with other methods.

5.4.4 Differential photometry

As the name suggests, differential photometry provides a differential measurement. That is to say that unlike the absolute photometric methods described above, the results

of differential photometry are relative to other reference sources in the same imaging frames. If the photometry of the reference objects is known in the absolute sense then differential photometry gives the photometry of the relevant object on the corresponding standard system as,

$$m = m_{\rm ref} - 2.5\log_{10} C_{\rm obj} + 2.5\log_{10} C_{\rm ref}, \tag{5.4.4}$$

where the subscript 'ref' differentiates the reference object from the target.

In this case absolute photometry was chosen over differential photometry, since there is no large optical catalogue of potential reference objects covering the entire VIKING field. The Deep Near Infrared Survey of the southern sky [DENIS; 148], provides a catalogue of bright I-band stars covering most of VIKING's southern field, but the majority of bright DENIS stars are saturated in the NTT images. The SDSS provides optical i- and z-band data over most of VIKING north, but lack of a similar data in VIKING south prevents the consistent use of differential photometry for every candidate. It does however provide a means to 'sanity check' results via an independent methodology.

The use of SDSS reference stars in the differential photometry results in target photometry on the SDSS AB system. To place the results on the NTT Vega system two corrections must be made to the reference stars prior to employing equation 5.4.4. The first is to apply the offset from the AB to the Vega system (Appendix C). The second is the conversion of SDSS to NTT magnitudes via a colour transformation equation. Figure 5.8 plots synthetic photometry of stars of spectral type O to M from the BPGS atlas [2] in the $i_{\rm NTT} - i$ versus $i - z$ plane (where i and z are the SDSS passbands). The straight line fit to the data relates NTT to SDSS stellar magnitudes as $i_{\rm NTT} = i - 0.4\,(i - z)$.

Figure 5.9 plots the absolute photometry versus differential photometry for candidates in the VIKING ngc field for which bight SDSS stars exist in the same images. The results are in good agreement, with residuals about the diagonal (solid line) of $\Delta_{\rm rms} = 0.15$ in either direction, which is comparable to the sky-background level as measured in the NTT imaging frames and depicted by the error bars. There is perhaps a slight systematic trend towards fainter magnitudes in the absolute photometry, which is probably partly related to the intermittent brightening of the $i_{\rm NTT}$ zero-point associated with two successive measurements of the same standard on night 3 of follow-up imaging (section 5.4.1). To test this theory the seven objects observed on night 3 are plotted as filled squares and the dashed line shows the best straight line fit assuming a slope of 1. The absolute photometry of these objects is indeed shifted towards fainter magnitudes, bringing the fit away from the diagonal, as would be expected for an overly bright estimate of the zero-point. The correction required to place night 3 magnitudes back onto the diagonal is $\Delta\, p \simeq 0.1$. A revised zero-point is re-calculated by omitting the offending standard star observations from the zero-point calculation and the result is shown in table 5.2, a correction $\Delta\, p \simeq 0.05$ is implied, less than the correction required.

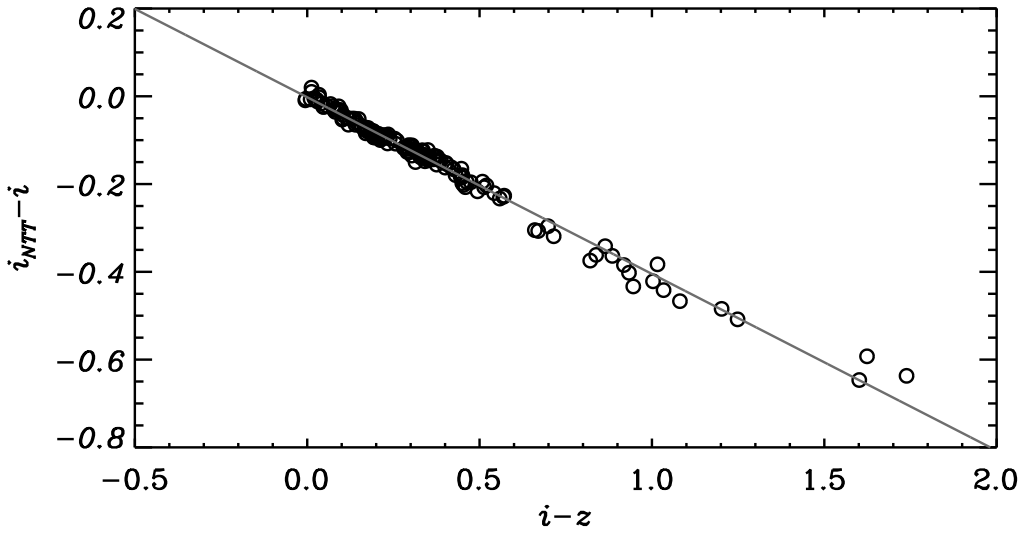


Figure 5.8: Straight line fit to stars in the BPGS spectroscopic atlas in the $i_{NTT}\,iz$ plane. The equation of the straight line is used to place bright SDSS stars of known absolute magnitude onto the NTT system in order to facilitate differential photometry. The fitting was performed by the `IDL` routine `LINFIT`.

The source of the remaining discrepancy is unknown but like the zero-point determination the aperture correction does not enter into the derivation of differential magnitudes and is thus a likely culprit. One could perhaps correct for this offset as implied by the fit, but since the overall agreement was found from the residuals about the diagonal, to be comparable to the uncertainties introduced by the sky background, this would seem unnecessary.

A similar analysis of the $z_{\rm NTT}$-band observations is not possible since only 2 of the $z_{\rm NTT}$ imaged candidates lie in regions of SDSS coverage. The magnitude offsets between the absolute and differential methods for these two sources are $\Delta\, z_{\rm NTT} = 0.34$ and $\Delta\, z_{\rm NTT} = 0.05$, the former being undetected above the background. The aim of $z_{\rm NTT}$ imaging is to acquire an independent verification of the VISTA photometry. Uncertainties in the aperture corrections amounting to a few tenths of a magnitude will not alter the confirmation of a significant drop in flux blueward of the VISTA Y-band given two independent measurements.

5.5 Results from imaging follow-up

A table of results is presented in the appendix B. This section concentrates on a subset of just six candidates which were undetected in $i_{\rm NTT}$ observations reaching 3σ depths of > 24 mag. A subset of table B.1 is re-produced for these candidates below (table 5.3). The remaining 38 candidates are visible in $i_{\rm NTT}$ and were consequently dropped from the candidate list. Here, $i_{\rm NTT}$ and $z_{\rm NTT}$ magnitudes are given in the asinh form, with

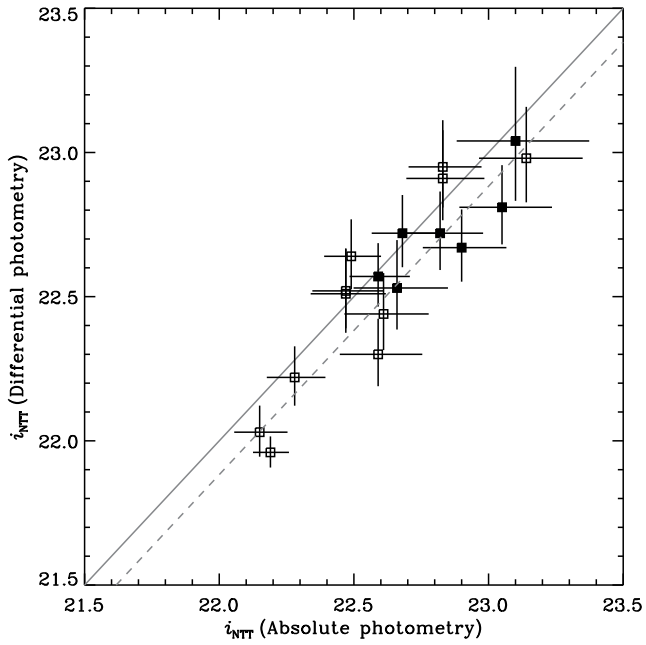


Figure 5.9: Comparison between the absolute and differential photometry of quasar candidates in regions of SDSS coverage. Bright SDSS stars in NTT imaging frames have been used to calibrate differential photometry. The results are in good agreement, with a residual about the diagonal (solid line) of 0.15. The best fit to the data (dashed line) reveals a slight offset towards brighter differential photometry. The offset is small $\Delta i_{\rm NTT} = 0.08$ and likely results from uncertainty in the aperture correction. Error bars show 1σ sky-noise about the measured magnitudes. A description of the fitting procedure is given by Williams et al. [3].

the softening parameters b equivalent to 25.1 in 24.1 in logarithmic magnitudes respectively.

Name	Z	δZ	Y	δY	J	δJ	$i_{\rm NTT}$	$\delta i_{\rm NTT}$	$z_{\rm NTT}$	$\delta z_{\rm NTT}$	*MCS*	*pStar*	*pGalaxy*	*pNoise*	P_q
#31	22.80	3.26	21.02	8.33	20.48	7.02	24.20	2.18	22.42	3.68	-0.924	1.000	0.000	0.000	1.47e-02
#32	23.38	1.70	20.85	10.43	20.18	9.41	24.89	0.64	24.70	-1.48	0.833	0.486	0.486	0.027	9.96e-01
sgp0109-3047	24.00	0.13	20.64	8.36	20.33	5.28	24.82	1.43	22.77	6.98	-	-	-	-	9.99e-01
sgp0305-3150	23.59	0.62	20.23	10.47	19.68	9.64	24.87	0.85	21.73	15.24	-	-	-	-	1.00e+00
sgp2218-3154	22.39	4.85	21.33	5.39	20.35	8.66	24.94	0.44	-	-	0.587	0.500	0.500	0.002	3.60e-05
sgp2348-3054	24.32	-0.37	20.44	11.55	20.30	8.54	25.61	-2.23	22.31	11.35	1.012	0.500	0.500	0.002	1.00e+00

Table 5.3: Photometric imaging results from the NTT for a subset of six objects which were undetected in deep $i_{\rm NTT}$-band imaging. For reference VISTA ZYJ magnitudes are also included. Photometric errors for each passband are given in columns headed with a leading 'δ' and are quoted as signal-to-noise, where the noise is the sky-background level. Detections with δY or δJ $\gtrsim$ 7.0 were selected for photometric follow-up by the techniques described in this thesis provided they also comply with the morphological constraints described in section 4.0.3. For reference the morphological attributes *MCS*, *pStar*, *pGalaxy* and *pNoise* and the probabilistic rank P_q of each source is also supplied. Hyphens indicate values that are unavailable for reasons discussed in the text. A full table of photometric imaging results in available in appendix B.

Although all candidates underwent probabilistic ranking prior to follow-up, the full list of targets is a result of work done by various members of the VIKING working group employing a number of independent approaches, so it is beneficial to look briefly at each source with reference to the selection strategy applied in this thesis. Each of the remaining candidates is afforded some brief discussion below. In most cases candidates are referred to via their equatorial coordinates (J 2000) prefixed by the relevant field e.g. ngc or sgp. In a two cases, follow-up work is on-going and the candidates concerned are referred to via the relevant row number table B.1.

sgp0109-3047 received a probabilistic rank $P_q = 1.0$. During the period of selection and up to the time of writing the relevant photometry and observation records are absent from the VSA. sgp0109-3047 was selected from the CASU catalogues via a pipeline developed independently to the work presented here. As such this candidate was not selected via the approach taken in this thesis. However based on P_q it is clearly a strong candidate and it is interesting to determine whether it would have been selected via the approach taken here.

Since sgp0109-3047 does not as yet appear in the VSA, band-merged morphological statistics such as *MCS* and *pclass* are not available. Independent passband level classifiers are given by CASU, which suggest a 'point-like' morphology in all passbands. Upon band merging, the confidence in these classifications will increase and it is safe to assume that the candidate with retain its point-like properties when it is finally ingested by the VSA.

As well morphological filtering, candidates were dropped if detected with a signal-to-noise below the 7σ level in either Y or J. With reference to table 5.3, sgp0109-3047 was detected with $(S/N)_Y = 8.4$ and $(S/N)_J = 5.3$ and thus, had it been present in the VSA, it would have been discarded prior to probabilistic ranking.

Perhaps these constraints are a little too conservative. With reference to figure 4.10, most of what is achieved in asking for a $S/N \geq 7\sigma$ detection in both Y and J is also achieved by asking for a $\geq 7\sigma$ detection in Y alone. In applying this more inclusive constraint, sgp0109-3047 would be put forward to the ranking stage. This is a valuable lesson learnt but the fact sgp0109-3047 did in reality make it to the follow-up stage, demonstrates the benefit of converging towards a common candidate list via a number of independent methods, as was the approach of the VIKING working group.

sgp0305-3150 presents a similar story to that of sgp0109-3047. Ranked with $P_q = 1.0$ it is another high priority candidate, which was absent from the VSA at the time of writing. It is designated as point-like in the CASU passband level catalogues and would therefore have passed the morphological constraints imposed in section 4.10. This source would also have passed the signal-to-noise constraints having been detected with $(S/N)_Y = 10.5$ and $(S/N)_J = 9.6$.

sgp2348-3054 was ranked with $P_q = 1.0$. The morphological characteristics of this

object are as follows; $MCS = 1.0$, $pgalaxy = 0.5$, $pstar = 0.5$ and $pnoise = 0.0$. It was detected with $(S/N)_Y = 11.5$ and $(S/N)_J = 8.5$ and it safely meets the selection constraints employed in this thesis.

sgp2218-3154 is a low priority candidate of $P_q = 3.6 \times 10^{-5}$. It has VIKING ZYJ photometry only marginally consistent with $z \geq 6.5$ quasars and would not normally have been observed but for a nightly period over which most higher priority candidates were not visible. Nevertheless it was undetected in i_{NTT} observations down to a 3σ depth of 24.3 mag and in light of these results it qualifies as a high priority candidate. The morphological properties of the source are point-like, $MCS = 0.6$, $pgalaxy = 0.5$, $pstar = 0.5$, $pnoise = 0.0$, and it was detected with $(S/N)_Y = 5.4$ and $(S/N)_J = 8.7$.

#32 is another highly ranked candidate with $P_q = 1.0$. Morphological attributes are consistent with a point source; $MCS = 0.8$, $pgalaxy = 0.5$, $pstar = 0.5$, $pnoise = 0.0$ and the detection significance is high; $(S/N)_Y = 10.4$ and $(S/N)_J = 9.4$. As such, the candidate safely meets the selection requirements laid down in this thesis.

#31 also meets the selection requirements having been ranked with $P_q = 1.5 \times 10^{-2}$ making it a modestly significant candidate. Morphological attributes are; $MCS = -0.9$, $pgalaxy = 0.0$, $pstar = 1.0$, $pnoise = 0.0$ and signal-to-noise ratios are $(S/N)_Y = 8.3$ and $(S/N)_J = 7.0$.

With the exception of sgp2218-3154, each of the above mentioned objects was further observed in z_{NTT}. The aim of these observations was to obtain independent confirmation of the VIKING Z-band photometry. As such these observations were made to a similar but in most cases a fainter limit than the VIKING observations.

It is difficult to compare z_{NTT}- with Z-band photometry directly, since the z_{NTT} filter possesses a red tail which bridges the Lyα transition until $z \simeq 7.2$. Conversely the VISTA Z-band possesses a step-like drop in sensitivity at its red extent, becoming insensitive to Lyα photons at $z \simeq 6.9$. Therefore the most straight forward approach is to consider the z_{NTT} YJ colour plane as shown in figure 5.10, where the colour scale gives P_q as derived from VIKING photometry. Error bars are computed from the sky-background level about the measured photometry and up-arrows give lower limits on the colours. The quasar and stellar loci are also plotted and by comparison the candidate colours are consistent with those of high-z quasars in the interval $6.6 \lesssim z \lesssim 7.0$.

This interval is perhaps better illustrated in the i_{NTT}Y z_{NTT} J colour plane in figure 5.11 panel (b), where the quasar tracks become redder in $Y - J$. If candidates are significantly fainter than the detection limit in i_{NTT} then they would fall along the quasar tracks between $z \simeq 6.6$ and $z \simeq 7.0$. Figure 5.11 panel (a) is further suggestive of the high-z nature of these candidates. Each of these candidates was put forward for spectroscopic observation.

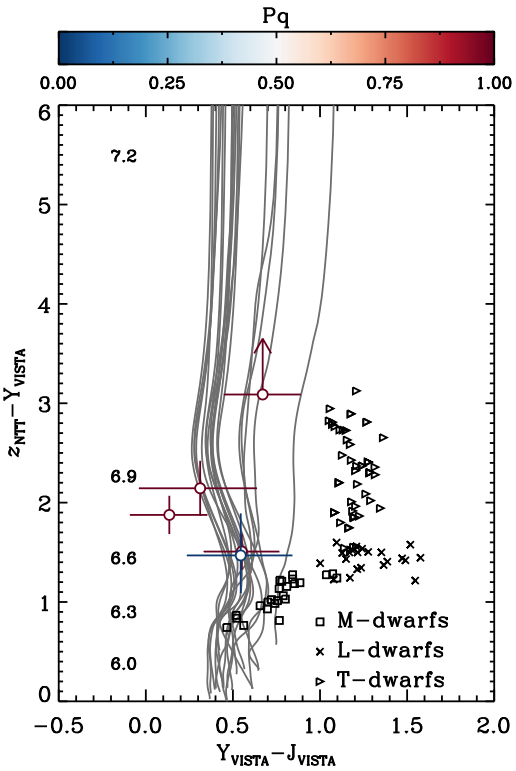
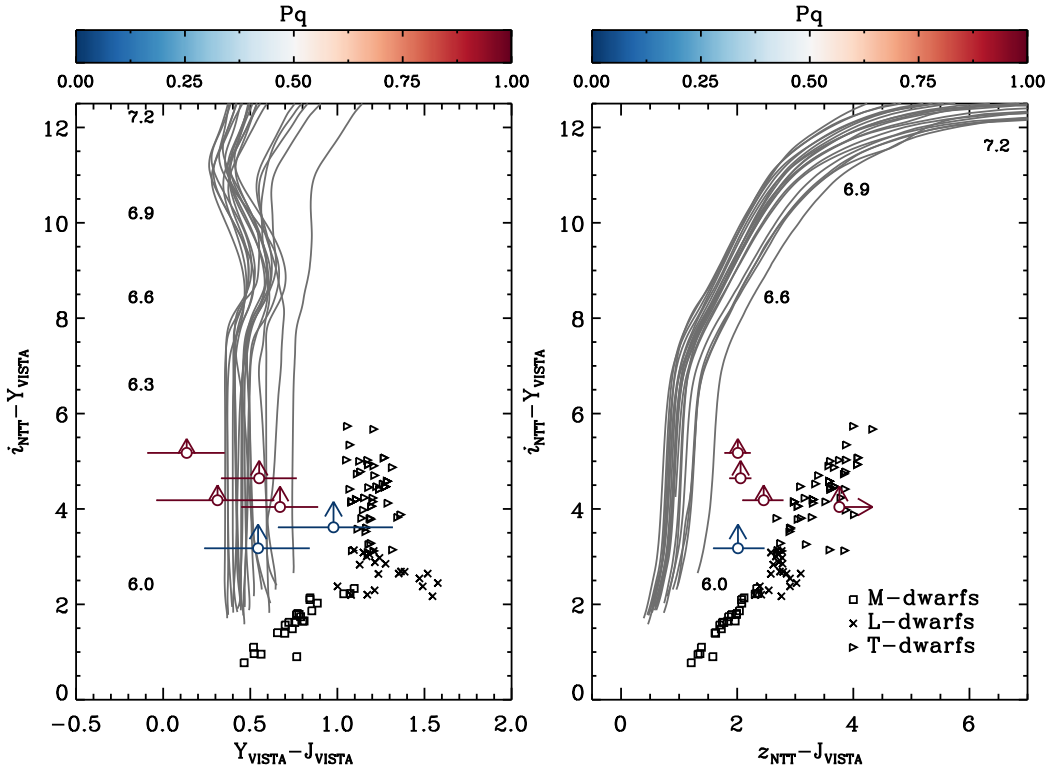


Figure 5.10: Five of the six candidates remaining after NTT imaging, (one candidate was not observed in z_{NTT} and is therefore omitted). Quasar candidates are plotted in the i_{NTT} z_{NTT} YJ plane with error-bars showing 1σ sky noise about the measured magnitudes, arrows represent lower limits. Each point is colour coded according to the probabilistic rank, P_q of the relevant candidate as derived from VIKING photometry. For comparison quasar redshift evolution tracks and cool-star synthetic photometry is also shown. All candidates are consistent with being $z \geq 6.5$ quasars.



(a) (b)

Figure 5.11: Panel (a): Remaining candidates after NTT imaging in the $i_{\rm NTT}$ YJ plane. Error bars show 1σ sky-noise about the measured magnitudes with arrows representing lower limits. For comparison quasar redshift evolution tracks and cool-star synthetic photometry is also shown. All remaining candidates are consistent with being high-z quasars. Panel (b): Candidates measured in both $i_{\rm NTT}$ and $z_{\rm NTT}$ are plotted in the $i_{\rm NTT}$Y versus $z_{\rm NTT}$J plane. The distribution of possible redshifts becomes clearer in this colour plane as the quasar locus begins to turn over in $z_{\rm NTT} - {\rm J}$.

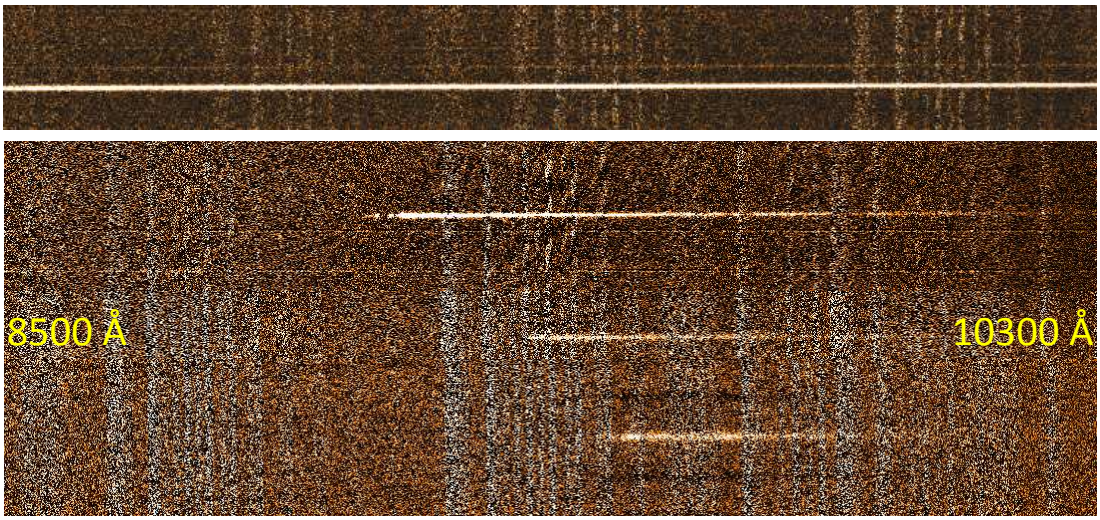


Figure 5.12: From top to bottom, 2-dimensional spectra of high-z quasar candidates sgp2218-3154, sgp0305-3150, sgp0109-3047 and sgp2348-3054. Approximate wavelengths at the extrema are given by the labels increasing left to right. In the lower three spectra a break in continuum flux is evident at wavelengths in excess of 9120 Å indicating three high-z objects. In the upper most spectrum close inspection shows a continuum a few pixels above the obvious bright source, that runs the length of the image, indicating a low-z origin. (credit: Dr. B. P. Venemans)

5.6 First spectra

Spectroscopic follow-up is currently on going and at the time of writing sgp0109-3047, sgp0305-3150, sgp2348-3054 and sgp2218-3154 have been observed on ESO's Focal reducer and low dispersion spectrograph II [FORS2; 149]. With the initial aim being the confirmation of these objects as quasars or otherwise, low signal-to-noise spectra were obtained in relatively bad weather conditions. On target integration time was $\sim$ 40 minutes in $\sim$ 1.5 arcsec seeing. The reduction and calibration of these spectra is work completed independently of this thesis, but for completeness the spectra are shown in 2- and 1-dimensional format in figure 5.12 and 5.13 respectively. Detailed discussion and analysis of these results will be given by Venemans et al. [150] in a forthcoming paper.

In figure 5.12, 2-dimensional spectra are shown from top to bottom for sgp2218-3154, sgp0305-3150, sgp0109-3047 and sgp2348-3054 respectively. Wavelength increases from left to right with the approximate wavelengths at the extrema being 8000 Å- 10000 Åas labelled in the figure. Spectra of light collected from various objects in the field of view of these observations (i.e. centered on the slit) are visible running left to right on each image with residuals from subtracted monochromatic sky lines running in the perpendicular direction.

The upper most spectrum is that of sgp2218-3154, the bright source in the image is a star $\sim$ 5 arcsec from the candidate. The continuum of sgp2218-3154 is visible just above background running along the centre of the image a few pixels above the bright

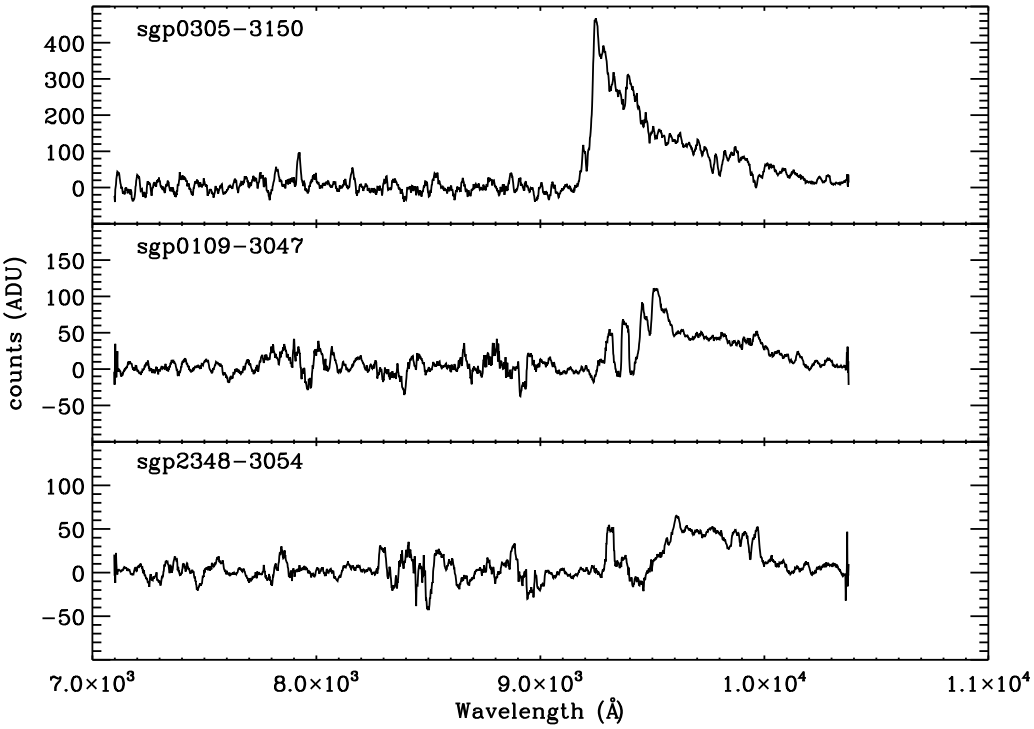


Figure 5.13: Extracted 1-dimensional spectra of sgp0305-3150, sgp0109-3047 and sgp2348-3054 confirming them unequivocally as quasars. (credit Dr. B. P. Venemans)

source. On close inspection the continuum extends over short to long wavelengths. If this were a $z \geq 6.5$ source one would expect to see a sharp drop in flux at the Lyα frequency, which would occur longward of $\lambda = 9120$ Å. This source can therefore be ruled out as a $z \geq 6.5$ quasar with high confidence.

For all other spectra a drop in continuum flux is apparent above $\lambda = 9120$ Å, indicating the presence of a high-z object. Figure 5.13 presents 1-dimensional spectra of these objects confirming that they are indeed $z \geq 6.5$ quasars. Preliminary analysis places them at $z = 6.5$, $z = 6.7$ and $z = 6.9$ [150], giving them the second, third and fourth highest redshifts of quasars discovered to date.

5.7 Quasar space density

The discovery of three of the four most distant quasars known represents a significant science highlight in the first year of VISTA science operations. Each of these three quasars will contribute widely to our understanding of the high-z Universe, serving a invaluable cosmological probes. This juncture would therefore seem like a fitting place to conclude, leaving detailed science on e.g. cosmic reionization and black hole growth to a time when high-resolution, high signal-to-noise spectra are available.

There are however some straight forward, if highly tentative, inferences that can be made with the new information supplied by these discoveries. In section 3.2.1 the space density of $6.5 \leq z \leq 7.5$ quasars was calculated via integration of the QLF for three different rates of evolution. Given that the search in the sgp is now complete, encompassing a total area of 180 deg^2, what can one infer about the evolution of quasar space density.

Section 3.2.1 considered a single evolutionary behaviour in which the QLF was subject to pure density evolution. Three rates of evolution were discussed; non-evolving, whereby the rate of evolution remains constant between $6.0 \leq z \leq 7.5$, a 'baseline' evolution corresponding to a factor 2 decline in density over each $\Delta z = 0.5$ interval in redshift and a more extreme evolution corresponding to a factor 3 decline over the same range.

Figure 5.15 plots the number of quasars brighter than a given J-band magnitude as predicted by each of the three models. The filled square represents the state of knowledge offered by the discovery of an extremely luminous $J = 19.3$ quasar at $z = 7.085$ [37] in the UKIDSS-LAS. The filled triangle shows the space density inferred from the three $J \leq 20.3$ VIKING quasars. The level of uncertainty on each data point is calculated assuming that quasars follow a Poissonian random distribution over the sky. The correct error calculation was the subject of some interesting discussion during the examination of this thesis. For the benefit of both the reader and the author the next few pages are updated to summarise the key points.

The Poisson distribution materialises from the binomial distribution in the limiting case where the probability of a single event is small but the number of trials is large providing opportunity to observe a number of successes. The probability of observing N successes over a given interval (time, frequency, length, area etc.) in this limit is

$$P(N) = \frac{e^{-\chi}\chi^N}{N!}, \tag{5.7.1}$$

where χ is the expectation value. If χ is large $P(N)$ becomes more symmetric and approaches a Gaussian distribution. This is the reason for the common use of the standard deviation of the Poisson distribution, $\sigma_{\text{Poiss}} = \sqrt{N}$, to approximate our level of uncertainty in a Poisson process as the width or the spread of a Gaussian distribution. When the number of successes over the interval of interest is small the Poisson distribution is asymmetric and this approach becomes inaccurate.

The problem with the $\sqrt{N}$ approximation becomes obvious when we consider applying it *incorrectly* as a means to quantify the spread in the sampling distribution relevant to the UKIDSS quasar search, which as mentioned has found one $z \geq 6.5$ quasar in ~ 2600 deg^2. In this case one would quote the best estimate $\hat{\chi}$ along with the 1σ Gaus-

sian error bar as $\frac{1\pm1}{2600}\,\mathrm{deg}^{-2}$. Taking this argument to the 3σ limit we have $\frac{1\pm3}{2600}\,\mathrm{deg}^{-2}$. So here, the incorrect use of Gaussian statistics manifests itself by allowing for the possibility of a negative count rate; a wholely unphysical prediction.

It is more relevant in this case to consider an asymmetric Poissonian sampling distribution. Such a distribution would peak at $\hat{\chi}$, but χ would be more likely found on one side of the peak than the other (see Figure 5.14). The use of error bars would be inappropriate here, since they are implicitly derived from symmetrical distributions.

In general there are two approaches better suited to this problem; that taken by the frequentist is founded in sampling theory and constructs an interval which includes the true value of the parameter of interest with a probability greater than or equal to a specified level, this is known as a confidence interval.

For example a frequentist might construct a 68% confidence interval for χ, such that with an infinite number of repeated samples, the true value would lie within this interval at least 68% of the time. The probability that χ lies within the confidence interval defined by any one of these samples is either 0 or 1 and thus the value of χ is fixed while the confidence interval, which depends on the random sample, is random.

For a sample $\mathbf{s} = \{s_i\}$ of N independent and identically distributed observations from a Poisson distribution with unknown expectation value χ, a confidence interval with endpoints $A(\mathbf{s})$ and $B(\mathbf{s})$ has the property that

$$P\left[A(\mathbf{s}) < \chi \leq B(\mathbf{s})\right] \geq \gamma \tag{5.7.2}$$

where $0 \leq \gamma \leq 1$ and $A(\mathbf{s})$ and $B(\mathbf{s})$ are commonly defined such that the probability masses in either tail of the distribution are equal.

The second approach to the problem, taken by Bayesians, is to consider the posterior probability distribution $P(\chi \,|\, \mathbf{s})$. Given any sample one can construct an interval that will have a specified probability of containing χ. This is known as a credible interval and is constructed such that

$$P\left[\, a(\mathbf{s}) < \chi \leq b(\mathbf{s}) \,|\, \mathbf{s} \,\right] = \gamma. \tag{5.7.3}$$

In contrast to equation 5.7.2, it is now $\mathbf{s}$ that is fixed while χ is distributed as $P(\chi \,|\, \mathbf{s})$.

Confidence intervals are described by [151]. The tools required to construct a credible interval were introduced to in Chapter 4 so it seems fitting to pursue this approach here.

The posterior probability distribution $P(\chi \,|\, \mathbf{s})$ is related to the prior probability distribution $P(\chi)$ and the likelihood function $P(\mathbf{s} \,|\, \chi)$ by

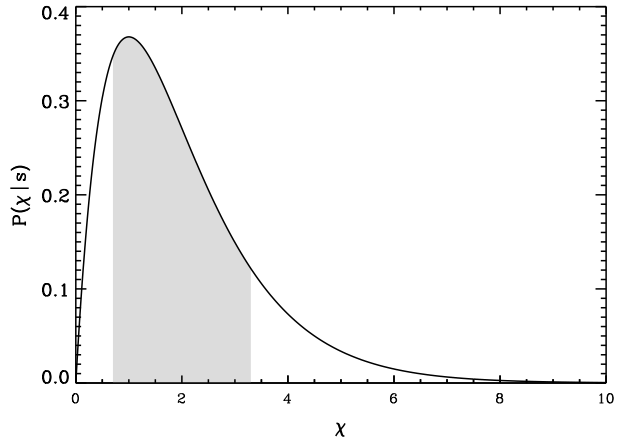


Figure 5.14: The probability of the expectation value χ, given one Poissonian distributed success in the context of the UKIDSS quasar search. The shaded region shows the 68% credible interval for the value of χ.

$$P\left(\chi \,|\, \mathbf{s}\right) \propto P\left(\chi\right) P(\mathbf{s} \,|\, \chi). \tag{5.7.4}$$

In the simplest case the prior is uniform and non-informative, $P\left(\chi\right) = 1$, and the posterior probability is completely defined by the Poisson likelihood function,

$$P(\mathbf{s}|\chi) = \prod_{i=1}^{N} \frac{\chi^{s_i}}{s_i!}\, e^{-\chi}, \tag{5.7.5}$$

which for the UKIDSS result is given by $P(s_1 = 1\,|\,\chi) = \chi\, e^{-\chi}$ and is shown in figure 5.14. In general to find a 68% credible interval the task is to find a and b such that

$$P\left(a < \chi \leq b \,|\, \mathbf{s}\right) = \frac{\int_a^b P\left(\mathbf{s} \,|\, \chi\right) d\chi}{\int_{-\infty}^{+\infty} P\left(\mathbf{s} \,|\, \chi\right) d\chi} = 0.68. \tag{5.7.6}$$

In this case, the posterior probability is a Poisson distribution so it is straight forward to find the interval $(\,a, b\,]$ via numerical integration. In the case of the UKIDSS quasar search this gives the credible interval shown by the shaded region in Figure 5.14 and the credible intervals relevant for both the VIKING and UKIDSS searches are shown over-plotted on surface density points in Figure 5.15.

The total area searched in the sgp is $180\,\mathrm{deg}^2$ yielding 3, $z \geq 6.5$ quasars down to a limiting magnitude of $J = 20.4$. The baseline number count model predicts 1.5 quasars at this limit. Taking this as the expectation value, one can use equation 5.7 to estimate the probability of three successes as 12.6 per cent. The non-evolving and the rapidly evolving scenarios predict 4.0 and 0.6 quasars respectively, returning likelihoods of 19.5 and 2.0 per cent.

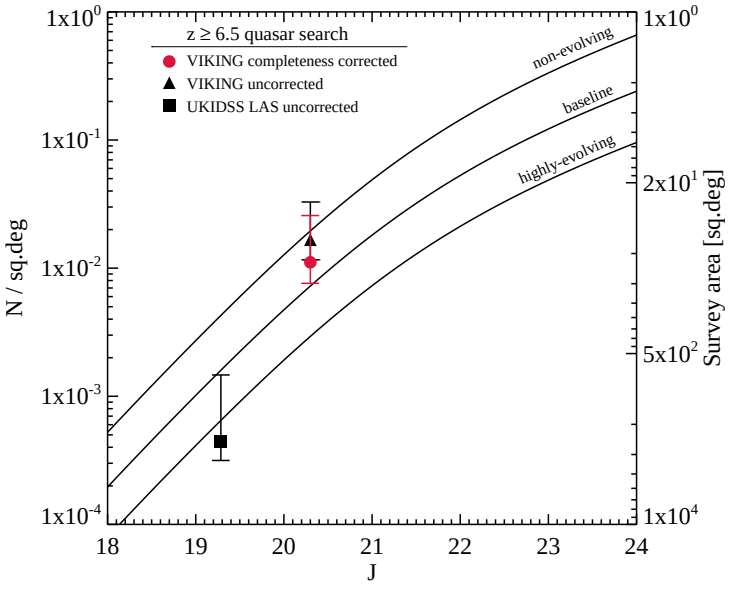


Figure 5.15: Predicted number density of $6.5 \leq z \leq 7.5$ quasars based on the QLF of [4, see section 3.2.1 for details]. Curves show a non-evolving, baseline and highly evolving decline in space density beyond $z = 6$. Space densities as indicated by results from the UKIDSS-LAS, VIKING and VIKING assuming a 67 per cent completeness are shown with Poissonian error bars.

At this stage however the completeness function of the quasar search is unclear, since it combines a number of different selection methods. At some stage this will need to be dealt with in detail by attempting to recover simulated objects planted in the VIKING imaging frames with use of the CASU source extractor [similar work has been undertaken by e.g. 4]. This is a non-trivial exercise and one which will warrant consideration when VIKING has covered substantially more area allowing for robust attempts to constrain quasar space density and ultimately the $z \geq 6.5$ QLF.

In section 4.4 an attempt to approximate the completeness function by modelling the signal-to-noise distribution of catalogued sources in the VIDEO xmm-3 field was made. The calculation combined the completeness implied by the morphological, signal-to-noise and probabilistic ranking constraints used to select quasars in this thesis. Two of the three quasars discovered in the sgp meet these selection requirements and the mean completeness over the redshift range $6.5 \leq z \leq 7.5$ and the magnitude interval $18.0 \leq J \leq 20.8$ is 67.0 per cent.

Given this selection function the non-evolving, baseline and rapidly-evolving scenarios predict 2.7, 1.01 and 0.41 quasars respectively. The likelihoods of finding 2 quasars given these predictions are 24.5, 18.6 and 5.6 per cent, marginally favouring the non-evolving model. From inspection of figure 5.15 however, both the VIKING and UKIDSS results are also consistent with the baseline model, i.e. no change in the rate of space

density evolution from z = 6.

With the small number statistics uncertainties in the QLF dominate over the number count predictions. At z $\sim$ 6 the QLF is well constrained at bright magnitudes ($M^*_{1450} < -25$) [4], but the QLF beyond z $\simeq$ 6.4 is completely unknown and it is difficult to place a limit in the level of confidence on the rate or behaviour of the evolution in quasar space density at these redshifts. Clearly this calculation will benefit greatly from better statistics. As expected VIKING is showing clear signs that these statistics will become available within the next 5 years of observations.

VISTA and VHS will confirm whether UKIDSS results, which favour a rapid rate of density evolution beyond z = 6, are real. If VHS results are consistent with UKIDSS but VIKING is finding a higher number of fainter quasars then this would suggest an interesting evolution in M^*_{1450}.

5.8 Summary

NTT optical imaging follow-up of 44 quasar candidates put forward by various members of the VIKING group have been reduced and corresponding photometry measured. This process is discussed in detail. Since z $\gtrsim$ 6.5 quasars are expected to be dark shortward of the $z_{\rm NTT}$ band, objects with clear detections in the $i_{\rm NTT}$ band are discarded, six objects remain and are shown to have $i_{\rm NTT}$, $z_{\rm NTT}$, Z, Y and J band photometry consistent with z $\geq$ 6.5 quasars. Spectroscopic identification of these candidates is ongoing. At the time of writing all four objects identified in the sgp region have been observed. Three of these four objects are high-z quasars with preliminary redshifts of z = 6.5, 6.7, 6.9 respectively.

A Poisson likelihood test shows that after taking the completeness function into account and considering only those quasars with properties consistent with the selection strategy employed in this thesis, that the rate of space density evolution in z $\geq$ 6.5 quasars marginally favours models which do not evolve significantly from z = 6. However, discovery rates from both VIKING and UKIDSS are also consistent with the rate of space density evolution seen a z = 6. This will need to be clarified with better completeness function estimates and further statistics.

The three quasars co-discovered here, represent a significant science highlight in the first year of the VISTA public surveys and will remain crucial probes of the high-z universe in the next decade.

CHAPTER 6

Summary conclusions & future work

The importance of high-z quasars as cosmological probes drives the effort to discover them at ever higher redshifts and hence ever earlier times. The VISTA Kilo-degree Infrared Galaxy Survey (VIKING) is expected to contain a significant sample of $z \gtrsim 6.5$ quasars. Such objects are observed less than 0.85 Gyr after the Big Bang, a period in which the Universe was undergoing or coming to the end of a series of great changes.

In particular it was a time that marked the luminous stage in the rise of the first population of massive black holes. These objects are the engines of quasars, consuming mass rapidly and radiating rest-mass energy. Explaining the existence and rapid growth indicated by masses of order a billion times solar at such an early stage, poses a significant theoretical challenge. The discovery of quasars at ever higher redshifts continues to place tight constraints on models of massive black hole formation and growth.

The growth of super-massive black holes and their massive galaxy hosts probably occurred synchronously in the early Universe. The most massive galaxies at high-z are extremely rare and are most likely to be found by the fact that they harbour super-massive black holes, which power luminous quasars. Quasar hosts offer excellent follow-up targets at radio and sub-mm wavelengths to probe star formation history in the earliest galaxies as well as co-eval black hole growth.

The build up of stellar material in the first galaxies led to the reionisation of cosmic hydrogen (and eventually helium). Absorption properties of the spectra of the highest redshift quasars indicate that the Universe or at least localised regions of it were less ionised in the past. The study of cosmic reionisation sets a benchmark for the formation of the first non-linear cosmic structures. Quasars provide the most straight forward means to obtain high signal-to-noise evidence of this epoch in cosmic history.

This thesis describes the isolation, follow-up and confirmation of the first $z \gtrsim 6.5$ quasars from VIKING. The remainder of this chapter summarises the work undertaken

and some routes for useful research to be carried out in the near future.

Chapter 2

The first full depth VIKING catalogue was made available to the VIKING working group in April 2011. Before commencing the quasar search it was felt necessary to implement some basic quality control; four distinct checks were considered.

First, as a broad indication of the quality and accuracy of the photometric measurements within the catalogue, re-measurements of point sources within overlapping regions between imaging frames and tiles were compared. The residuals in these measurements were found to be broadly conformative with the survey requirements, an indication of the high quality of the archived photometry.

Second, the lack of a large sample of $z \gtrsim 6.5$ quasar spectra necessitates the need to model the photometric properties of high-z quasars in the VISTA passbands. The modelling must be carried out relative to a standard photometric system and since all VISTA data is defined with reference to the standard star Vega it would seem natural to anchor any modelled photometry to this system too.

However, VISTA data is zero-pointed via a set of colour equations with reference to a set of 2MASS standard stars. The basic assumption being that the colour equations over the JHKs wavelengths, can be linearly extrapolated to cover the Z and Y passbands. This was never expected to be a robust means to place VISTA Z and Y photometry onto the Vega system but rather it was intended to provide a good first order approximation.

Offsets from the Vega system in the VIKING photometry were investigated via the relative distribution of red and blue stars in the various SDSS and VIKING colour spaces. Significant offsets from the Vega system were found in the Z and Y bands ($\Delta_Z = -0.116 \pm 0.008$, $\Delta_Y = -0.094 \pm 0.007$), and these were applied to all subsequent Vega zeroed synthetic photometry in order to place it in line with VISTA .

Third, VIKING shares a region of overlap with the VIDEO survey in the near-infrared as well as the CFHTLS Deep 1 field in the optical. Both surveys provide deep observations of $\sim 1\,\mathrm{deg}^2$ patch of the VIKING field and offer the chance to study VIKING sources with precision nine-band photometry. The big advantage of this data set is that it allows accurate star galaxy separation in optical-near-infrared colour spaces. The VIKING catalogue was matched to each of the VIDEO and CFHTLS surveys ready to be called into use in subsequent chapters.

In doing this, significant offsets from the VIKING photometry were noted in the VIDEO photometry. In plotting the offset as a function of VIKING photometry it was further noted that offsets appeared to be constant as a function of brightness. Straight lines were fit to these distributions and offsets were applied to bring the VIDEO photometry in line with the VIKING photometry.

Fourth, two classes of artifact, which cause contamination of the VIKING catalogue and in particular quasar candidate lists, exist in the VISTA images. Both are found in and around the vicinity of bright (and in many cases saturated) stars. The first of these appendages is the result of microscopic imperfections in VISTA's primary and secondary mirrors, causing diffraction spikes that can either produce false detections or compromise the photometry of real sources. The second class of artifacts arise from reflections within the filter and/or detector interfaces, causing roughly circular ghost images. Like the diffraction spikes, these ghosts can produce either false sources or compromise the photometry of real sources.

The provision of a recipe to remove objects influenced by the presence of diffraction spikes is provided via an empirical fit to the distribution of radial distances from bright star intensity peaks as a function of stellar brightness. A similar recipe is provided to remove objects lying in bright star ghosts by considering the optical properties of the VISTA telescope and camera.

As discussed in chapter 2, these quality control checks and quality insurance procedures are beneficial to the quasar search. However they are also beneficial in range science applications. The importance of the broad overview of the quality of the VIKING photometry given here is clear for any science application. There are also a number of reasons for one to want to know the extent of the offset of VIKING photometry from the Vega system, a plausible example might be in the application to accurate photometric redshifts. Image artifacts in the near vicinity of bright stars will effect all data samples drawn from the VIKING catalogue. The recipes presented in this thesis for their flagging or removal are tools that could be employed widely by users of the VIKING data.

Chapter 3

In this chapter the use of the VIKING ZYJ colour plane in the selection of $6.5 \lesssim z \lesssim 7.5$ quasars is demonstrated. With the first VIKING catalogue covering only a small fraction ($\sim 200\,\mathrm{deg}^2$) of the overall footprint, it is argued that in order to conduct the fairest non-biased quasar search it is necessary to investigate contamination of quasar colour space via use of a combination of initial data and detailed modelling.

A combination of synthetic photometry and number count models are employed to populate a mock catalogue of the cool star component of the VIKING survey in the ZYJ passbands. The catalogue incorporates realistic Gaussian scatter as determined from the sky background measurements in VIKING imaging frames. The signal-to-noise distribution for detections in the real catalogue is modelled in order to mimic the results of the source extraction algorithm. This is then imposed on the simulated data. The model is compared to $\sim 1\,\mathrm{deg}^2$ of imaging in the VIKING-VIDEO-CFHTLS overlap region using photometry from the complementary *gi* JKs colours as a robust

star-galaxy separator. The model and data are found to be in good agreement and the mock catalogue provides a means to investigate cool-star contamination in quasar selection strategies in lieu of a large sample of real VIKING data.

Similarly a number count model of the $6.5 \lesssim z \lesssim 7.5$ quasar population is constructed via integration of the latest determination of the quasar luminosity function. The findings imply that the current data set provides a fair possibility of containing a $z \gtrsim 6.5$ quasar at the VIKING J-band 10σ limiting magnitude.

Chapter 4

In chapter 4 contamination of the ZYJ colour plane by cool-stars and moderate redshift galaxies is investigated. To begin with the VIKING catalogues are queried for objects that have colours consistent with $z \gtrsim 6.5$ quasars.

Consecutively a set 'broad-brush' constraints are placed on the catalogue, which include a detection significance threshold, the removal of detections with significant error bits and the removal of objects effected by diffraction or ghosting around bright stars. Secondly a highly complete sample of VIKING galaxies classified via their optical-near-infrared colours are used to set constraints on the morphological characteristics of faint point sources in the VIKING catalogues. These constraints are used to remove objects with extended morphologies from the quasar search. A set of colour constraints are then employed to limit contamination to a reasonable number for follow-up

For comparison a second method of selection is then investigated. The mock VIKING catalogue is employed as a distinct hypothesis in as Bayesian model comparison; the opposing hypothesis being a modelled catalogue high-z quasars. Members of a test set of simulated quasars are then ranked on the probability that they are quasars over stars.

The completeness functions of each of these methods are compared and it is found that a search which uses the Bayesian ranking method is on average $\sim$ 10 per cent more complete than a simple colour selection technique. On the basis of this outcome a catalogue of high-z quasar candidates is defined via probabilistic ranking. These candidates are then visually inspected for quality control and those that pass this final test are put forward for follow-up imaging.

In previous works that have used Bayesian ranking for quasar selection, a stellar model is constrained via a maximum likelihood fit to the real data. In the case of new surveys such as VIKING this is not possible with the limited data expected from initial releases. Therefore this methodology is relevant to a number recently commencing or forthcoming rare object searches in new imaging surveys such as PanSTARRS and VST-KIDS.

Investigation into the morphological attributes of sources in this chapter will allow future VIKING users to determine the completeness and contamination of point-source

samples by galaxies or vice-versa

Chapter 5

Chapter 5 describes the imaging and subsequent reduction of a set of high-z quasar candidates put forward by various members of the VIKING group for follow-up on the ESO's New Technology telescope in June 2011. After the reduction of $i_{\rm NTT}$ and $z_{\rm NTT}$ band imaging, just six objects remained consistent with being z $\gtrsim$ 6.5 quasars. These objects are currently undergoing a program of spectroscopic follow-up. To date all four objects found in the VIKING sgp field have had spectroscopic follow-up. Three of these objects have been confirmed as high-z quasars with redshifts of $z = 6.5, 6.7, 6.9$. We await spectroscopic confirmation of a further two candidates in the ngc field.

By comparison with expectations from the quasar luminosity function and employing Poisson likelihood test, the space density of quasars implied by these discoveries is consistent with no evolution in space density between z $\sim$ 6-7, but only marginally more so than a space density that declines at the same rate as seen between z $\sim$ 3-6. Further discoveries will help to place better constraints on the evolution of the QLF beyond z $\sim$ 6.

Future work

So what next for these new discoveries? A number of followup studies along the same lines as those outlined in the introduction of this thesis are possible. Here, attention is given to the most likely follow-up studies to be undertaken by the VIKING working group within the next six months.

Although it is difficult to put strong quantitative and global constraints on reionisation with tests on high-z quasars (see chapter 1). They do nevertheless provide useful suggestive trends on the physical processes taking place at these epochs. Gunn-Peterson like tests on line of sight IGM neutral fractions are out of the question with these three quasars. They are discovered at a time when the IGM volume averaged neutral fraction was too large and serve as upper limits in this respect only.

However some useful tests have recently been carried on the most distant quasar yet discovered at z = 7.085 [ULAS J1120+0641, hereafter J1120; 37]. High signal-to-noise spectra were obtained on the VLT FORS2 and Gemini near-infrared spectrograph (GNIRS) on Gemini North and show clear differences to all quasars yet discovered at z $\sim$ 6.

The first of these is the size of the quasar ionisation zone, which by convention is inferred from the radius at which the measured fraction of light transmitted blueward of the Lyα transition falls to 0.1. Quasar ionisation zone sizes at z $\sim$ 6, are on average $R_{\rm ion} = (7.4 - 8.0\,[z - 6])$ Mpc [31], the quasar at z $\simeq$ 7 has an ionisation zone

$R_{\rm ion} = 2.1 \pm 0.1\,{\rm Mpc}$.

As discussed in chapter 1, there is large scatter and uncertainty in making these measurements, but if the three VIKING quasars also show such clear departures from expectations at $z \simeq 6$ it will at least confirm that the Universe continues to undergo an accelerated state of change between $6 \lesssim z \lesssim 7$.

A tantalising indication of significantly neutral gas in the vicinity of the J1120 is the likely detection of the Lyα damping wing [152]. The presence of a damping wing may be understood by considering the finite width of the Lyα transition. The probability distribution for an emitted frequency is a highly peaked function about $\lambda_{{\rm Ly}\alpha} = 1216\,$Å; a convolution of a Lorentzian profile with intrinsic width reflecting the finite lifetime in an excited state, and a Gaussian reflecting Doppler broadening. The resulting line has a Voigt profile. In a highly neutral IGM a small fraction of Lyα emission will be absorbed at frequencies slightly longward and slightly shortward of $\lambda_{{\rm Ly}\alpha}$ and one should not expect a sharp discontinuity at $\lambda_{{\rm Ly}\alpha}$ but rather a damped wing.

Shortward of Lyα, absorption due to the dampening effect is difficult to distinguish from absorption in the GP trough region, but longward of Lyα the damping would become clear. J1120 is the first quasar to show signs of this effect consistent with a neutral fraction $x_{\rm HI} > 0.1$ [153] in the surrounding IGM. The only other source bright enough for the damping wing to have been detected is GRB050904 a GRB at $z \sim 6.3$, which duly faded after a few days. Similar detections in the VIKING quasars would be indicative of an epoch in which the neutral fraction of even the most biased regions of the Universe are still partly neutral.

It is a natural step to mirror these successes with similar high signal-to-noise observations of the three VIKING quasars for comparison. This may not be straight forward however, J1120 has an absolute magnitude $M_{1450,{\rm AB}} = -26.6$, significantly brighter than the typical magnitude of the VIKING quasars, $M_{1450,{\rm AB}} = -25.7$.

The chance to directly map the distribution of HI around quasars buried in a significantly neutral IGM was recently put forward by Cantalupo et al. [154]. The emissivity of the relatively low density IGM is dominated by two processes; radiative recombination of free electrons by protons and collisional excitation of HI via interaction with energetic free electrons. The former is favoured if there is a very high ionisation fraction. Provided there is a significant neutral gas content, the latter is a steeply increasing function of IGM temperature between $2 - 5 \times 10^4\,$K (figure 1 from [154]) and can increase the Lyα signal by several orders of magnitude where conditions are favourable.

These conditions are found in the ionisation fronts of high-z quasars. As the ionisation front expands within the IGM it will encounter neutral patches. Lyα collisional excitation can then take place in the transition zone between the ionisation front and neutral IGM, the subsequent emission giving a 2D "tomography" of the neutral patches of IGM encountering the ionisation front.

This signal will be easily detected with the reduced background and large collecting areas offered by forthcoming space borne projects such JWST. However if the signal, as expected, originates from regions of extended width a few arcmins on the sky, then high resolution spectroscopy from current facilities could detect individual neutral patches by integrating the signal over a fractions of the slit length.

In lieu of detectable emission from single neutral patches, the integral emission can be used to obtain better measurements of the scale of the ionisation front placing constraints on the IGM neutral fraction. Observations are likely to be carried out J1120 soon, it is perhaps best to wait for these results before attempting similar on the fainter VIKING quasars.

With the discovery of three more $z \gtrsim 6.5$ quasars to add to J1120, there is clearly a wealth of scientific discovery to come in the near future.

Concluding remarks

Quasars are the brightest non-transient objects in the high-z universe, far outshining stars and normal galaxies in all regions of the electromagnetic spectrum. Quasars are now consistently found at $z \sim 6$ corresponding to look back times of just a few hundred million years after the Big Bang. These objects represent the rise of the first population of super massive black holes and remarkable evidence of the ability to form evolved structures of several billion solar masses in the very early universe. Since their discovery and the detection of their cosmological redshifts, quasars have been used as cosmological probes in fields as far reaching as black hole growth and galaxy assembly, star formation and chemical enrichment and the evolution of cosmic hydrogen after the dark ages.

Now, a new generation of wide field quasar searches is beginning to build upon the recent success of most notably the SDSS and the CFHTLS. Prompted by technological advances in near infrared and optical imaging, and the pioneering work in the near infrared by UKIDSS in recent years, these surveys are motivating the discovery of quasars at unprecedented depths and redshifts. Work in this field will remain at the forefront of scientific exploration, aided by a wealth of data from forthcoming facilities.

VISTA is already providing deep complementary near-infrared data to UKIDSS and in the near future deep optical data sets will be available from the VST and the Pan-STARRS. In the more distant future dark energy related projects such as the LSST and Euclid will provide tens of quasars at $z \sim 10$ and hundreds at $6.5 \lesssim z \lesssim 8.5$

It is rare that a set of such exceptional state of the art facilities overlap to work independently towards a common science goal. With the ability to catalogue these objects in large numbers comes the opportunity to ask statistically interesting questions about the evolving Universe and serves as strong motivation for quasar searches and science

with the resulting samples. To optimise the science, these resources should be treated as complementary data sets and the synergy between them should be fully exploited.

I am pleased to have been able to contribute to the forthcoming wave of science expected from above mentioned projects. The discovery of three $z \geq 6.5$ quasars is a considerable science highlight in the first year of VISTA operations and places the VIKING high-z quasar search onto the best possible footing for the future.

APPENDIX A

Straight line fitting procedure

In fitting a 2-dimensional distribution to a specific model, the fitting procedure must take into account the error distribution in both dimensions. Where there are significant errors on both axes and perhaps also an intrinsic scatter, a familiar straight line linear regression is not the correct solution and can give biased results. This is illustrated in figure A.1, where figure 2.4 is reproduced with a further two fits using straight forward linear regression. In this case the short dashed green line shows the fit in which $J - K_s$ is assumed to take on values that are known to infinite precision (or at least to a significantly higher precision than in $Y - J$). The opposite scenario is shown by the long dashed green line. The optimised solution lies somewhere in between and is shown by the blue line as calculated in chapter 2.

A correct approach would deal with all the information available by incorporating the full error distribution in the problem. One way to do this is to maximise the likelihood function of the data. This is the approach taken in this thesis and in chapter 2 the method is used to locate the offset from the Vega system in the VISTA Z and Y bands from straight line fits to blue stars in the ZJKs and YJKs colour planes. The need for this analysis is discussed in chapter 2, here the general procedure is outlined.

Given a statistical model with m adjustable parameters a_j, $j = 0...m-1$, to describe the underlying distribution of an observed population of N data points (x_i, y_i), $i = 0...N-1$, the likelihood of a given set of parameter values given the observations is equivalent to the product of the probabilities that each data point should occur from the model distribution. Formally, this is given by,

$$\mathcal{L} = \prod_{i=0}^{N-1} P(x_i, y_i \mid a_0 ... a_{m-1}). \tag{A.0.1}$$

Here x_i and y_i are data points drawn from an underlying distribution $Y_i = f(X_i \mid a_0 ... a_{m-1})$ and are subject to measurement errors δx_i, δy_i, with variance σ^2_{xi}, σ^2_{yi}, such that $x_i = X_i + \delta x_i$ and $y_i = Y_i + \delta y_i$.

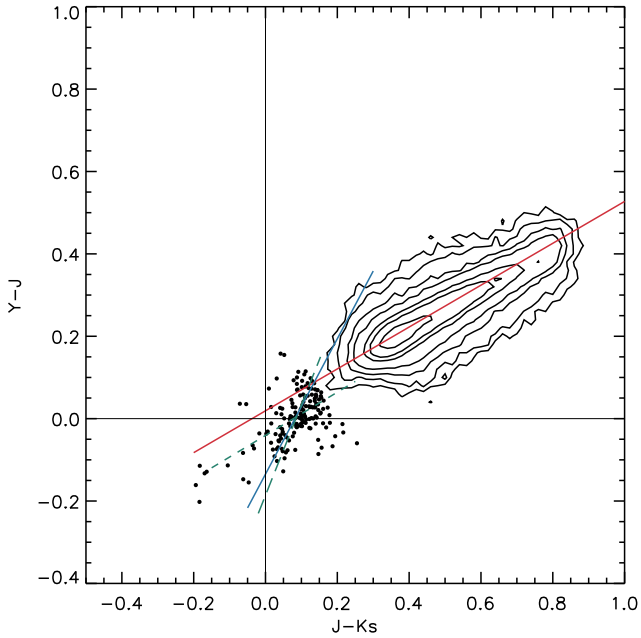


Figure A.1: As in figure 2.4 but with the addition of standard linear regression fits with errors in Y − J (dashed green line) and J − Ks (long dashed green line). The optimum solution can not be found by a standard linear regression and is instead given by a maximum likelihood fit with consideration of errors in both coordinates simultaneously. The optimum solution is shown by the blue line as calculated in chapter 2.

For independently Gaussian distributed errors,

$$\mathcal{L} \propto \prod_{\mathrm{i}=0}^{N-1} \left\{ \exp\left[-\frac{1}{2}\left(\frac{\delta x_{\mathrm{i}}^2}{\sigma_{x\mathrm{i}}^2} + \frac{\delta y_{\mathrm{i}}^2}{\sigma_{y\mathrm{i}}^2} \right) \right] \right\}. \tag{A.0.2}$$

The best fit model parameters occur where the likelihood takes its maximum value and the calculation then involves finding derivatives. In this case the fit is to a set of statistically independent points and the likelihood of drawing the entire data set from the model is the product of the individual likelihoods of drawing each data point from the model. Therefore it is easier to differentiate $\ln(\mathcal{L})$, which is the sum of individual log-likelihoods,

$$\ln(\mathcal{L}) \propto -\frac{1}{2} \sum_{\mathrm{i}=0}^{N-1} \frac{(x_{\mathrm{i}} - X_{\mathrm{i}})^2}{\sigma_{x\mathrm{i}}^2} + \frac{(y_{\mathrm{i}} - Y_{\mathrm{i}})^2}{\sigma_{y\mathrm{i}}^2}. \tag{A.0.3}$$

Since $\ln(\mathcal{L})$ is a monotonically decreasing function it achieves its minimum at the same point at $\mathcal{L}$ achieves its maximum, defined by the following differential constraints

$$\frac{\partial \ln(\mathcal{L})}{\partial X_{\mathrm{i}}} = \frac{(x_{\mathrm{i}} - X_{\mathrm{i}})}{\sigma_{x\mathrm{i}}^2} + \frac{\partial f}{\partial X_{\mathrm{i}}} \left[\frac{y_i - f(X_{\mathrm{i}} \,|\, a_0 \ldots a_{\mathrm{m}-1})}{\sigma_{y\mathrm{i}}^2} \right] = 0, \tag{A.0.4}$$

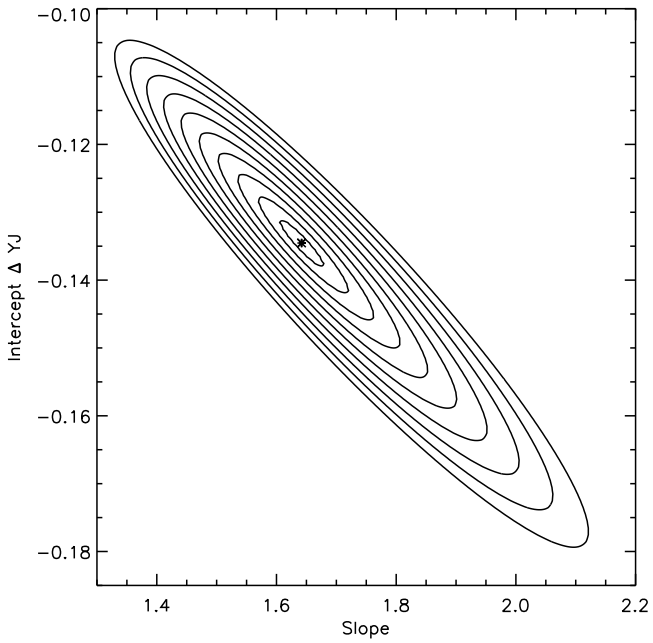


Figure A.2: The Log-likelihood contours corresponding to the maximum likelihood fit to the blue star locus shown in figure 2.4. Contours start at 1σ and are spaced in intervals of 1σ assuming a Gaussian distribution of scatter in the fitted parameters.

$$\frac{\partial \ln(\mathcal{L})}{\partial a_{\rm j}} = \sum_{\rm i=0}^{N-1} \frac{\partial f}{\partial a_{\rm j}} \left[\frac{y_{\rm i} - f(X_{\rm i} \mid a_0 ... a_{\rm m-1})}{\sigma^2_{y{\rm i}}} \right] = 0. \tag{A.0.5}$$

For a straight line model $Y_{\rm i} = a_0 X_{\rm i} + a_1$, equation A.0.4 reduces to

$$\frac{x_{\rm i} - X_{\rm i}}{\sigma^2_{x{\rm i}}} + \frac{a_0 (y_{\rm i} - a_0\, X_{\rm i} - a_1)}{\sigma^2_{x{\rm i}}} = 0, \tag{A.0.6}$$

which may be re-arranged to give

$$X_{\rm i} = \frac{a_0\, y_{\rm i} - a_0\, a_1 + x_{\rm i}\, \epsilon^2_{\rm i}}{a_0^2 + \epsilon^2_{\rm i}}, \tag{A.0.7}$$

where $\epsilon^2_{\rm i} = (\sigma_{y{\rm i}}/\sigma_{x{\rm i}})^2$. With expressions for $X_{\rm i}$ and $Y_{\rm i}$, equation A.0.3 can by minimised via a suitable optimisation routine. Here the algorithm of choice is the Nelder-Mead simplex [155] implemented by the Interactive Data Language (`IDL`) `Amoeba` routine.

Log-likelihood contours corresponding to the fit in figure 2.4 are shown in figure A.2.

APPENDIX B

Follow-up photometry

The results of the NTT follow-up photometry of the high-z quasar candidates discussed in chapter 5 are presented below. For reference VISTA ZYJ magnitudes are also included. The signal-to-noise of each detection is given in columns headed with a leading 'δ'. The noise measurements in the NTT photometry includes contributions from the sky background only. The noise measurements for VISTA photometry are calculated by CASU from the average sky background in the relevant imaging frames and includes contributions from the source, sky and inter-pixel correlations. Morphological attributes *MCS*, *pStar*, *pGalaxy* and *pNoise* and the probabilistic rank P_q based on VSA photometry are further supplied. Hyphens indicate values that are unavailable for reasons discussed in the chapter 5. Each source is referred to via its ID which is made up of it's equatorial sky coordinates (J 2000) pre-fixed by the VIKING field in which it is found e.g. ngc, sgp. For a few objects follow-up is ongoing and ID is replaced by the row number in the table below.

ID	Z	δZ	Y	δY	J	δJ	$i_{\rm NTT}$	$\delta i_{\rm NTT}$	$z_{\rm NTT}$	$\delta z_{\rm NTT}$	*MCS*	*pstar*	*pgalaxy*	*pnoise*	Pq
ngc1146+0145	21.47	3.75	20.14	12.82	19.54	8.56	22.56	7.14	-	-	0.521	0.993	0.003	0.003	3.00e-06
ngc1147+0126	21.74	2.91	20.31	11.02	19.53	8.58	22.44	7.83	-	-	-0.156	0.994	0.003	0.003	6.60e-06
ngc1149-0046	22.01	5.32	20.56	9.10	19.94	8.23	22.58	7.01	-	-	0.366	0.945	0.052	0.003	1.73e-08
ngc1219+0003	21.98	2.61	20.63	4.33	20.28	6.03	22.12	11.03	-	-	0.844	0.944	0.052	0.003	1.94e-05
ngc1211+0007	22.57	3.46	20.69	8.50	19.67	8.47	22.43	8.26	-	-	-1.141	0.994	0.003	0.003	3.00e-05
ngc1218+0106	21.66	8.92	20.46	12.17	19.84	9.58	22.22	10.04	-	-	-0.578	0.486	0.486	0.027	4.38e-20
ngc1418+0109	22.48	4.51	21.09	8.11	20.69	7.35	22.80	7.54	-	-	1.747	0.499	0.499	0.002	1.25e-06
ngc1419-0100	22.62	3.57	20.68	10.64	21.39	1.88	22.80	8.07	-	-	2.148	0.050	0.900	0.050	2.84e-01
ngc1431+0001	23.96	0.28	21.03	8.32	20.04	13.87	22.90	0.19	-	-	-1.317	0.994	0.003	0.003	3.26e-02
ngc1418-0018	21.99	6.68	20.48	12.76	19.96	7.01	22.43	10.35	-	-	1.001	0.945	0.052	0.003	3.55e-10
sgp2215-3145	20.92	19.61	20.71	9.57	20.74	6.03	21.03	45.81	-	-	0.082	0.945	0.052	0.003	3.49e-83
sgp2222-3129	22.93	2.18	21.08	6.23	20.69	7.48	23.07	6.87	-	-	1.153	0.499	0.499	0.002	4.43e-03
sgp0125-3016	22.90	1.16	20.45	7.35	19.59	9.61	22.24	13.72	-	-	-	-	-	-	1.34e-02
sgp0221-3139	22.22	2.64	20.39	9.99	19.86	12.01	22.04	16.13	-	-	0.637	0.997	0.003	0.000	5.48e-03
sgp0223-3208	22.17	5.22	20.61	9.26	20.77	6.28	22.29	15.37	-	-	2.068	0.003	0.997	0.000	1.34e-06
ngc1425-0054	22.32	5.36	20.55	12.31	19.93	10.43	22.17	16.25	-	-	0.921	0.945	0.052	0.003	8.19e-06
ngc1413+0139	22.22	5.82	20.88	9.18	20.57	5.22	23.06	5.75	-	-	1.223	0.052	0.945	0.003	3.22e-09
sgp0223-3144	21.93	3.50	20.60	8.22	19.82	12.48	23.16	5.05	-	-	1.325	0.997	0.003	0.000	3.10e-07
sgp0259-3100	23.33	1.12	20.67	7.94	20.16	6.07	22.12	12.31	-	-	-	-	-	-	9.27e-01
sgp0305-3400	22.92	1.22	20.35	9.72	19.54	10.28	22.12	13.09	-	-	-	-	-	-	1.25e-01
ngc1213+0200	22.06	6.30	20.62	10.20	19.97	9.10	23.03	5.26	-	-	0.714	1.000	0.000	0.000	8.38e-11
ngc1218+0051	22.04	2.49	20.74	3.90	20.08	7.28	22.60	8.25	-	-	0.957	0.997	0.003	0.000	2.00e-05
ngc1412+0128	22.30	5.40	20.82	9.72	20.41	6.06	22.86	7.93	-	-	1.323	0.945	0.052	0.003	1.05e-07

Continued overleaf

ID	Z	δZ	Y	δY	J	δJ	$i_{\rm NTT}$	$\delta i_{\rm NTT}$	$z_{\rm NTT}$	$\delta z_{\rm NTT}$	*MCS*	*pstar*	*pgalaxy*	*pnoise*	Pq
ngc1447+0113	21.97	4.98	20.80	8.00	20.05	7.83	22.64	10.09	-	-	2.751	0.000	1.000	0.001	2.11e-09
ngc1450-0023	22.36	5.02	20.96	8.41	21.01	2.74	23.00	7.11	-	-	-1.163	0.945	0.052	0.003	2.75e-07
ngc1414+0126	22.14	6.35	20.89	9.12	20.50	5.57	22.75	8.48	-	-	0.688	0.945	0.052	0.003	4.76e-11
ngc1434-0003	22.13	6.45	20.77	10.76	20.20	8.93	22.54	10.88	-	-	0.164	0.945	0.052	0.003	3.82e-11
sgp2358-3355	23.67	0.57	20.84	7.74	20.24	7.31	22.74	7.27	-	-	0.593	0.486	0.486	0.027	8.25e-01
ngc0914+0155	22.42	4.89	20.97	9.21	20.45	9.62	23.52	4.32	-	-	1.162	0.499	0.499	0.002	3.15e-07
ngc1151+0118	22.61	3.98	20.93	9.04	20.60	5.68	23.21	7.39	-	-	2.289	0.486	0.486	0.027	2.89e-04
#31	22.80	3.26	21.02	8.33	20.48	7.02	24.20	2.18	22.42	3.68	-0.924	1.000	0.000	0.000	1.47e-03
#32	23.38	1.70	20.85	10.43	20.18	9.41	24.89	0.64	24.70	-1.48	0.833	0.486	0.486	0.027	9.96e-01
ngc1429+0142	22.33	4.13	20.55	10.72	19.83	8.07	23.00	11.89	-	-	0.314	0.944	0.052	0.003	9.72e-05
ngc1450+0100	22.16	3.11	20.37	8.73	19.71	7.52	22.15	20.69	-	-	4.142	0.000	1.000	0.000	5.77e-04
ngc1451+0021	21.81	4.33	20.30	9.26	19.67	7.81	23.18	7.63	-	-	0.004	1.000	0.000	0.000	1.31e-06
sgp0109-3047	24.00	0.13	20.64	8.36	20.33	5.28	24.82	1.43	22.77	6.98	-	-	-	-	9.99e-01
sgp0305-3150	23.59	0.62	20.23	10.47	19.68	9.64	24.87	0.85	21.73	15.24	-	-	-	-	1.00e+00
sgp0322-3350	23.41	0.60	20.83	4.41	20.21	5.11	22.95	11.44	-	-	-	-	-	-	1.34e-02
sgp2214-3134	22.86	2.94	21.38	5.14	20.90	5.24	23.41	9.50	-	-	1.064	0.500	0.500	0.002	9.07e-05
sgp2218-3154	22.39	4.85	21.33	5.39	20.35	8.66	24.94	0.44	-	-	0.587	0.500	0.500	0.002	3.60e-05
sgp2223-3158	22.16	6.22	20.91	9.49	20.20	11.76	23.36	7.09	-	-	3.098	0.000	1.000	0.000	3.05e-12
sgp2226-3124	23.03	1.95	21.26	5.27	20.63	7.90	23.48	7.05	-	-	0.125	0.997	0.003	0.000	1.26e-03
sgp2229-3057	23.39	1.03	20.98	7.24	20.45	4.91	23.54	6.03	-	-	1.025	0.052	0.947	0.000	2.83e-01
sgp2348-3054	24.32	-0.37	20.44	11.55	20.30	8.54	25.61	-2.23	22.31	11.35	1.012	0.500	0.500	0.002	1.00e+00

APPENDIX C

Photometric offsets

VIDEO and CFHTLS data in use throughout this thesis has been transformed from the AB (absolute) system onto the Vega system. The AB system is defined to have a constant flux density over all wavelength intervals, the reference spectrum being flat in $f(\nu) \equiv f_\nu \equiv \text{const}\,[\text{ergs}\,\text{s}^{-1}\,\text{cm}^{-2}\,\text{Hz}^{-1}]$. The constant is defined such that $f_\nu\,\mathrm{d}\nu \equiv f_\lambda\,\mathrm{d}\lambda$ at the effective wavelength of the Johnson V-band $\lambda_V = 5480\,\text{Å}$.

Offsets from the VIDEO (VIRCAM) and CFHTLS (MegaCam) AB magnitudes to their Vega equivalents are given for each passband in table C.1. These are calculated by evaluating equation 3.1.1 for the SED of Vega and a flat $f_\nu = \text{const}$ spectrum over the relevant passbands, the difference in magnitudes giving the offset. For reference table C.1 also supplies the central wavelength of each filter.

Two MegaCam i_{CFH} filters are listed; due to a mechanical failure which led to the destruction of $i_{\text{CFH}}(1)$, a new filter $i_{\text{CFH}}(2)$ was commissioned in June 2007. $i_{\text{CFH}}(1)$ was operational on MegaCam for CFHTLS nights between mid-2003 to mid-2007, $i_{\text{CFH}}(2)$ became the new *i*-band filter until the survey end in early 2009. Reference to the i_{CFH}-band in this thesis corresponds to $i_{\text{CFH}}(1)$. Offsets in the SDSS and WFCAM filters can be found in [72].

Instrument	Filter	Central wavelength (Å)	Conversion AB to Vega
MegaCam	u_{CFH}	3743	-0.338
	g_{CFH}	4872	0.093
	r_{CFH}	6282	-0.164
	i_{CFH} (1)	7776	-0.389
	i_{CFH} (2)	-	-0.376
	z_{CFH}	11702	-0.551
VIRCAM	Z	8800	-0.512
	Y	10200	-0.600
	J	12500	-0.916
	H	16500	-1.360
	Ks	21500	-1.827

Table C.1: Offsets from the AB to the Vega system for the MegaCam and VIRCAM *ugriz* and ZYJHKs filter sets employed by the CFHTLS and VISTA surveys respectively.

References

[1] D. E. Vanden Berk, G. T. Richards, A. Bauer, M. A. Strauss, D. P. Schneider, T. M. Heckman, D. G. York, P. B. Hall, X. Fan, G. R. Knapp, S. F. Anderson, J. Annis, N. A. Bahcall, M. Bernardi, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, F. J. Castander, A. J. Connolly, J. H. Crocker, I. Csabai, M. Doi, D. Finkbeiner, S. Friedman, J. A. Frieman, M. Fukugita, J. E. Gunn, G. S. Hennessy, Ž. Ivezić, S. Kent, P. Z. Kunszt, D. Q. Lamb, R. F. Leger, D. C. Long, J. Loveday, R. H. Lupton, A. Meiksin, A. Merelli, J. A. Munn, H. J. Newberg, M. Newcomb, R. C. Nichol, R. Owen, J. R. Pier, A. Pope, C. M. Rockosi, D. J. Schlegel, W. A. Siegmund, S. Smee, Y. Snir, C. Stoughton, C. Stubbs, M. SubbaRao, A. S. Szalay, G. P. Szokoly, C. Tremonti, A. Uomoto, P. Waddell, B. Yanny, and W. Zheng. Composite Quasar Spectra from the Sloan Digital Sky Survey. *AJ*, 122:549–564, August 2001. doi: 10.1086/321167.

[2] J. E. Gunn and L. L. Stryker. Stellar spectrophotometric atlas, wavelengths from 3130 to 10800 Å. *ApJS*, 52:121–153, June 1983. doi: 10.1086/190861.

[3] M. J. Williams, M. Bureau, and M. Cappellari. The Tully-Fisher relations of early-type spiral and S0 galaxies. *MNRAS*, 409:1330–1346, December 2010. doi: 10.1111/j.1365-2966.2010.17406.x.

[4] C. J. Willott, P. Delorme, C. Reylé, L. Albert, J. Bergeron, D. Crampton, X. Delfosse, T. Forveille, J. B. Hutchings, R. J. McLure, A. Omont, and D. Schade. The Canada-France High-z Quasar Survey: Nine New Quasars and the Luminosity Function at Redshift 6. *AJ*, 139:906–918, March 2010. doi: 10.1088/0004-6256/139/3/906.

[5] B. Chen, C. Stoughton, J. A. Smith, A. Uomoto, J. R. Pier, B. Yanny, Ž. Ivezić, D. G. York, J. E. Anderson, J. Annis, J. Brinkmann, I. Csabai, M. Fukugita, R. Hindsley, R. Lupton, J. A. Munn, and SDSS Collaboration. Stellar Population Studies with the SDSS. I. The Vertical Distribution of Stars in the Milky Way. *ApJ*, 553:184–197, May 2001. doi: 10.1086/320647.

[6] M. Jurić, Ž. Ivezić, A. Brooks, R. H. Lupton, D. Schlegel, D. Finkbeiner, N. Padmanabhan, N. Bond, B. Sesar, C. M. Rockosi, G. R. Knapp, J. E. Gunn, T. Sumi,

D. P. Schneider, J. C. Barentine, H. J. Brewington, J. Brinkmann, M. Fukugita, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, E. H. Neilsen, Jr., A. Nitta, S. A. Snedden, and D. G. York. The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution. *ApJ*, 673:864–914, February 2008. doi: 10.1086/523619.

[7] J. A. Caballero, A. J. Burgasser, and R. Klement. Contamination by field late-M, L, and T dwarfs in deep surveys. *A&A*, 488:181–190, September 2008. doi: 10.1051/0004-6361:200809520.

[8] M. Stiavelli, S. G. Djorgovski, C. Pavlovsky, C. Scarlata, D. Stern, A. Mahabal, D. Thompson, M. Dickinson, N. Panagia, and G. Meylan. Evidence of Primordial Clustering around the QSO SDSS J1030+0524 at z=6.28. *ApJL*, 622:L1–L4, March 2005. doi: 10.1086/429406.

[9] Y. Shen, M. A. Strauss, M. Oguri, J. F. Hennawi, X. Fan, G. T. Richards, P. B. Hall, J. E. Gunn, D. P. Schneider, A. S. Szalay, A. R. Thakar, D. E. Vanden Berk, S. F. Anderson, N. A. Bahcall, A. J. Connolly, and G. R. Knapp. Clustering of High-Redshift ($z \geq 2.9$) Quasars from the Sloan Digital Sky Survey. *AJ*, 133:2222–2241, May 2007. doi: 10.1086/513517.

[10] R. Maiolino. Early metal enrichment in high-redshift quasars. In G. Giobbi, A. Tornambe, G. Raimondo, M. Limongi, L. A. Antonelli, N. Menci, & E. Brocato, editor, *American Institute of Physics Conference Series*, volume 1111 of *American Institute of Physics Conference Series*, pages 160–167, May 2009. doi: 10.1063/1.3141537.

[11] X. Fan, M. A. Strauss, R. H. Becker, R. L. White, J. E. Gunn, G. R. Knapp, G. T. Richards, D. P. Schneider, J. Brinkmann, and M. Fukugita. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z$\sim$6 Quasars. II. A Sample of 19 Quasars. *AJ*, 132:117–136, July 2006. doi: 10.1086/504836.

[12] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *ApJS*, 180:330–376, February 2009. doi: 10.1088/0067-0049/180/2/330.

[13] J. P. Emerson, W. J. Sutherland, A. M. McPherson, S. C. Craig, G. B. Dalton, and A. K. Ward. The Visible & Infrared Survey Telescope for Astronomy. *The Messenger*, 117:27–32, September 2004.

[14] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120:1579–1587, September 2000. doi: 10.1086/301513.

[15] J. A. Peacock. *Cosmological Physics*. Cambridge, UK: Cambridge University Press, January 1999.

[16] F. Hamann and G. Ferland. Elemental Abundances in Quasistellar Objects: Star Formation and Galactic Nuclear Evolution at High Redshifts. *ARA&A* , 37:487–531, 1999. doi: 10.1146/annurev.astro.37.1.487.

[17] A. C. Phillips. *Introduction to quantum mechanics*. Chichester, UK: John Wiley & Sons Ltd., January 2003.

[18] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18, February 2011. doi: 10.1088/0067-0049/192/2/18.

[19] J. S. B. Wyithe and A. Loeb. Reionization of Hydrogen and Helium by Early Stars and Quasars. *ApJ*, 586:693–708, April 2003. doi: 10.1086/367721.

[20] T. R. Choudhury and A. Ferrara. Experimental constraints on self-consistent reionization models. *MNRAS*, 361:577–594, August 2005. doi: 10.1111/j.1365-2966.2005.09196.x.

[21] J. S. Bolton and M. G. Haehnelt. The observed ionization rate of the intergalactic medium and the ionizing emissivity at z $\geq$ 5: evidence for a photon-starved and extended epoch of reionization. *MNRAS*, 382:325–341, November 2007. doi: 10.1111/j.1365-2966.2007.12372.x.

[22] S. R. Furlanetto, M. Zaldarriaga, and L. Hernquist. The Growth of H II Regions During Reionization. *ApJ*, 613:1–15, September 2004. doi: 10.1086/423025.

[23] I. T. Iliev, G. Mellema, U.-L. Pen, H. Merz, P. R. Shapiro, and M. A. Alvarez. Simulating cosmic reionization at large scales - I. The geometry of reionization. *MNRAS*, 369:1625–1638, July 2006. doi: 10.1111/j.1365-2966.2006.10502.x.

[24] T. R. Choudhury, M. G. Haehnelt, and J. Regan. Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Lyα forest data. *MNRAS*, 394:960–977, April 2009. doi: 10.1111/j.1365-2966.2008.14383.x.

[25] J. E. Gunn and B. A. Peterson. On the Density of Neutral Hydrogen in Intergalactic Space. *ApJ*, 142:1633–1641, November 1965. doi: 10.1086/148444.

[26] R. H. Becker, X. Fan, R. L. White, M. A. Strauss, V. K. Narayanan, R. H. Lupton, J. E. Gunn, J. Annis, N. A. Bahcall, J. Brinkmann, A. J. Connolly, I. Csabai, P. C. Czarapata, M. Doi, T. M. Heckman, G. S. Hennessy, Ž. Ivezić, G. R. Knapp, D. Q. Lamb, T. A. McKay, J. A. Munn, T. Nash, R. Nichol, J. R. Pier, G. T. Richards, D. P. Schneider, C. Stoughton, A. S. Szalay, A. R. Thakar, and D. G. York. Evidence for Reionization at z$\sim$6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar. *AJ*, 122:2850–2857, December 2001. doi: 10.1086/324231.

[27] R. L. White, R. H. Becker, X. Fan, and M. A. Strauss. Probing the Ionization State of the Universe at z $\geq$ 6. *AJ*, 126:1–14, July 2003. doi: 10.1086/375547.

[28] M. McQuinn, S. P. Oh, and C.-A. Faucher-Giguère. On Lyman-limit Systems and the Evolution of the Intergalactic Ionizing Background. *ApJ*, 743:82, December 2011. doi: 10.1088/0004-637X/743/1/82.

[29] A. Songaila and L. L. Cowie. Approaching Reionization: The Evolution of the Ly α Forest from z=4 to z=6. *AJ*, 123:2183–2196, May 2002. doi: 10.1086/340079.

[30] P. Paschos and M. L. Norman. A Statistical Analysis of Intergalactic Medium Transmission Approaching Reionization. *ApJ*, 631:59–84, September 2005. doi: 10.1086/431787.

[31] C. L. Carilli, R. Wang, X. Fan, F. Walter, J. Kurk, D. Riechers, J. Wagg, J. Hennawi, L. Jiang, K. M. Menten, F. Bertoldi, M. A. Strauss, and P. Cox. Ionization Near Zones Associated with Quasars at z $\sim$ 6. *ApJ*, 714:834–839, May 2010. doi: 10.1088/0004-637X/714/1/834.

[32] A. Maselli, S. Gallerani, A. Ferrara, and T. R. Choudhury. On the size of HII regions around high-redshift quasars. *MNRAS*, 376:L34–L38, March 2007. doi: 10.1111/j.1745-3933.2007.00283.x.

[33] I. D. McGreer, A. Mesinger, and X. Fan. The first (nearly) model-independent constraint on the neutral hydrogen fraction at z $\sim$ 6. *MNRAS*, 415:3237–3246, August 2011. doi: 10.1111/j.1365-2966.2011.18935.x.

[34] A. Mesinger. Was reionization complete by z $\sim$ 5-6? *MNRAS*, 407:1328–1337, September 2010. doi: 10.1111/j.1365-2966.2010.16995.x.

[35] M. McQuinn, A. Lidz, M. Zaldarriaga, L. Hernquist, and S. Dutta. Probing the neutral fraction of the IGM with GRBs during the epoch of reionization. *MNRAS*, 388:1101–1110, August 2008. doi: 10.1111/j.1365-2966.2008.13271.x.

[36] B. E. Robertson, R. S. Ellis, J. S. Dunlop, R. J. McLure, and D. P. Stark. Early star-forming galaxies and the reionization of the Universe. *Nature*, 468:49–55, November 2010. doi: 10.1038/nature09527.

[37] D. J. Mortlock, S. J. Warren, B. P. Venemans, M. Patel, P. C. Hewett, R. G. McMahon, C. Simpson, and T. et al. Theuns. A luminous quasar at a redshift of z = 7.085. *Nature*, 474:616–619, June 2011. doi: 10.1038/nature10159.

[38] M. F. Morales and J. S. B. Wyithe. Reionization and Cosmology with 21-cm Fluctuations. *ARA&A* , 48:127–171, September 2010. doi: 10.1146/annurev-astro-081309-130936.

[39] J. S. B. Wyithe. Redshifted 21-cm observations of high-redshift quasar proximity zones. *MNRAS*, 387:469–480, June 2008. doi: 10.1111/j.1365-2966.2008.13267.x.

[40] M. J. Rees. Emission from the nuclei of nearby galaxies - Evidence for massive black holes. In E. M. Berkhuijsen & R. Wielebinski, editor, *Structure and Properties of Nearby Galaxies*, volume 77 of *IAU Symposium*, pages 237–242, 1978.

[41] S. G. Djorgovski, M. Volonteri, V. Springel, V. Bromm, and G. Meylan. The Origins and the Early Evolution of Quasars and Supermassive Black Holes. In H. Kleinert, R. T. Jantzen, & R. Ruffini, editor, *The Eleventh Marcel Grossmann Meeting On Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories*, pages 340–367, September 2008. doi: 10.1142/9789812834300_0018.

[42] C. Degraf, T. Di Matteo, and V. Springel. Faint-end quasar luminosity functions from cosmological hydrodynamic simulations. *MNRAS*, 402:1927–1936, March 2010. doi: 10.1111/j.1365-2966.2009.16018.x.

[43] C. J. Willott, L. Albert, D. Arzoumanian, J. Bergeron, D. Crampton, P. Delorme, J. B. Hutchings, A. Omont, C. Reylé, and D. Schade. Eddington-limited Accretion and the Black Hole Mass Function at Redshift 6. *AJ*, 140:546–560, August 2010b. doi: 10.1088/0004-6256/140/2/546.

[44] F. Shankar. The demography of supermassive black holes: Growing monsters at the heart of galaxies. *New. Ast. Rev*, 53:57–77, April 2009. doi: 10.1016/j.newar.2009.07.006.

[45] C. J. Willott. No Evidence of Obscured, Accreting Black Holes in Most z = 6 Star-forming Galaxies. *ApJL*, 742:L8, November 2011. doi: 10.1088/2041-8205/742/1/L8.

[46] M. Volonteri. Formation of supermassive black holes. *A&A. Rev*, 18:279–315, July 2010. doi: 10.1007/s00159-010-0029-x.

[47] L. Ferrarese and H. Ford. Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research. *Space Sci. Rev.*, 116:523–624, February 2005. doi: 10.1007/s11214-005-3947-6.

[48] C. J. Lonsdale, D. Farrah, and H. E. Smith. Ultraluminous Infrared Galaxies. In Mason, J. W., editor, *Astrophysics Update 2*, page 285. Springer Verlag, 2006.

[49] L. Jiang, X. Fan, M. Vestergaard, J. D. Kurk, F. Walter, B. C. Kelly, and M. A. Strauss. Gemini Near-Infrared Spectroscopy of Luminous z $\sim$ 6 Quasars: Chemical Abundances, Black Hole Masses, and Mg II Absorption. *AJ*, 134:1150, September 2007. doi: 10.1086/520811.

[50] Y. Juarez, R. Maiolino, R. Mujica, M. Pedani, S. Marinoni, T. Nagao, A. Marconi, and E. Oliva. The metallicity of the most distant quasars. *A&A*, 494:L25–L28, February 2009. doi: 10.1051/0004-6361:200811415.

[51] C. Warner, F. Hamann, and M. Dietrich. A Relation between Supermassive Black Hole Mass and Quasar Metallicity? *ApJ*, 596:72–84, October 2003. doi: 10.1086/377710.

[52] D. P. Stark, R. S. Ellis, K. Chiu, M. Ouchi, and A. Bunker. Keck spectroscopy of faint 3 $\leq$ z $\leq$ 7 Lyman break galaxies - I. New constraints on cosmic reionization from the luminosity and redshift-dependent fraction of Lyman α emission. *MNRAS*, 408:1628–1648, November 2010. doi: 10.1111/j.1365-2966.2010.17227.x.

[53] R. Wang, C. L. Carilli, J. Wagg, F. Bertoldi, F. Walter, K. M. Menten, A. Omont, P. Cox, M. A. Strauss, X. Fan, L. Jiang, and D. P. Schneider. Thermal Emission from Warm Dust in the Most Distant Quasars. *ApJ*, 687:848–858, November 2008. doi: 10.1086/591076.

[54] R. Wang, C. L. Carilli, R. Neri, D. A. Riechers, J. Wagg, F. Walter, F. Bertoldi, K. M. Menten, A. Omont, P. Cox, and X. Fan. Molecular Gas in z $\sim$ 6 Quasar Host Galaxies. *ApJ*, 714:699–712, May 2010. doi: 10.1088/0004-637X/714/1/699.

[55] L. Jiang, X. Fan, W. N. Brandt, C. L. Carilli, E. Egami, D. C. Hines, J. D. Kurk, G. T. Richards, Y. Shen, M. A. Strauss, M. Vestergaard, and F. Walter. Dust-free quasars in the early Universe. *Nature*, 464:380–383, March 2010. doi: 10.1038/nature08877.

[56] H. Hao, M. Elvis, F. Civano, and A. Lawrence. Hot-dust-poor Quasars in Mid-infrared and Optically Selected Samples. *ApJ*, 733:108, June 2011. doi: 10.1088/0004-637X/733/2/108.

[57] C. Hazard, M. B. Mackey, and A. J. Shimmins. Investigation of the Radio Source 3C 273 By The Method of Lunar Occultations. *Nature*, 197:1037–1039, March 1963. doi: 10.1038/1971037a0.

[58] M. Ryle and A. Sandage. The Optical Identification of Three New Radio Objects of the 3c 48 Class. *ApJ*, 139:419, January 1964. doi: 10.1086/147770.

[59] R. H. Becker, R. L. White, and D. J. Helfand. The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. *ApJ*, 450:559, September 1995. doi: 10.1086/176166.

[60] M. Schmidt and R. F. Green. Quasar evolution derived from the Palomar bright quasar survey and other complete quasar surveys. *ApJ*, 269:352–374, June 1983. doi: 10.1086/161048.

[61] B. J. Boyle, T. Shanks, and B. A. Peterson. The evolution of optically selected QSOs. II. *MNRAS*, 235:935–948, December 1988.

[62] D. P. Schneider, M. Schmidt, and J. E. Gunn. The Palomar Transit GRISM Survey. In *American Astronomical Society Meeting Abstracts*, volume 25 of *Bulletin of the American Astronomical Society*, page 1402, December 1993.

[63] P. C. Hewett, C. B. Foltz, and F. H. Chaffee. The large bright quasar survey. 6: Quasar catalog and survey parameters. *AJ*, 109:1498–1521, April 1995. doi: 10.1086/117380.

[64] S. J. Warren, P. C. Hewett, and P. S. Osmer. A wide-field multicolor survey for high-redshift quasars, Z greater than or equal to 2.2. 3: The luminosity function. *ApJ*, 421:412–433, February 1994. doi: 10.1086/173660.

[65] M. Irwin, R. G. McMahon, and C. Hazard. APM optical surveys for high redshift quasars. In D. Crampton, editor, *The Space Distribution of Quasars*, volume 21 of *Astronomical Society of the Pacific Conference Series*, pages 117–126, 1991.

[66] J. D. Kennefick, R. R. de Carvalho, S. G. Djorgovski, M. M. Wilber, E. S. Dickson, N. Weir, U. Fayyad, and J. Roden. The Discovery of Five Quasars at z $>$ 4 Using the Second Palomar Sky Survey. *AJ*, 110:78, July 1995. doi: 10.1086/117498.

[67] S. J. Warren, P. C. Hewett, M. J. Irwin, R. G. McMahon, and M. T. Bridgeland. First observation of a quasar with a redshift of 4. *Nature*, 325:131–133, January 1987. doi: 10.1038/325131a0.

[68] C. J. Willott, X. Delfosse, T. Forveille, P. Delorme, and S. D. J. Gwyn. First Results from the Canada-France High-z Quasar Survey: Constraints on the z = 6 Quasar Luminosity Function and the Quasar Contribution to Reionization. *ApJ*, 633:630–637, November 2005. doi: 10.1086/462408.

[69] N. Kaiser, H. Aussel, and B. E. Burke. Pan-STARRS: A Large Synoptic Survey Telescope Array. In J. A. Tyson & S. Wolff, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4836 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 154, 2002.

[70] A. Lawrence, S. J. Warren, O. Almaini, A. C. Edge, N. C. Hambly, R. F. Jameson, P. Lucas, and M. Casali. The UKIRT Infrared Deep Sky Survey (UKIDSS). *MNRAS*, 379:1599–1617, August 2007. doi: 10.1111/j.1365-2966.2007.12040.x.

[71] S. Warren and P. Hewett. WFCAM, UKIDSS, and z = 7 Quasars. In N. Metcalfe and T. Shanks, editors, *A New Era in Cosmology*, volume 283 of *Astronomical Society of the Pacific Conference Series*, page 369, 2002.

[72] P. C. Hewett, S. J. Warren, S. K. Leggett, and S. T. Hodgkin. The UKIRT Infrared Deep Sky Survey ZY JHK photometric system: passbands and synthetic colours. *MNRAS*, 367:454–468, April 2006. doi: 10.1111/j.1365-2966.2005.09969.x.

[73] B. P. Venemans, R. G. McMahon, S. J. Warren, E. A. Gonzalez-Solares, P. C. Hewett, D. J. Mortlock, S. Dye, and R. G. Sharp. The discovery of the first luminous z $\sim$ 6 quasar in the UKIDSS Large Area Survey. *MNRAS*, 376:L76–L80, March 2007. doi: 10.1111/j.1745-3933.2007.00290.x.

[74] D. J. Mortlock, M. Patel, S. J. Warren, et al. Discovery of a redshift 6.13 quasar in the UKIRT infrared deep sky survey. *A&A*, 505:97–104, October 2009. doi: 10.1051/0004-6361/200811161.

[75] M. Patel et al. *MNRAS*, in preparation, 2011.

[76] B. P. Venemans et al. *MNRAS*, in preparation, 2011.

[77] M. F. Skrutskie, R. M. Cutri, R. Stiening, M. D. Weinberg, S. Schneider, J. M. Carpenter, C. Beichman, R. Capps, T. Chester, J. Elias, J. Huchra, J. Liebert, C. Lonsdale, D. G. Monet, S. Price, P. Seitzer, T. Jarrett, J. D. Kirkpatrick, J. E. Gizis, E. Howard, T. Evans, J. Fowler, L. Fullmer, R. Hurt, R. Light, E. L. Kopan, K. A. Marsh, H. L. McCallon, R. Tam, S. Van Dyk, and S. Wheelock. The Two Micron All Sky Survey (2MASS). *AJ*, 131:1163–1183, February 2006. doi: 10.1086/498708.

[78] M. Arnaboldi, M. J. Neeser, L. C. Parker, P. Rosati, M. Lombardi, J. P. Dietrich, and W. Hummel. ESO Public Surveys with the VST and VISTA. *The Messenger*, 127:28, March 2007.

[79] S. P. Driver, D. T. Hill, L. S. Kelvin, A. S. G. Robotham, J. Liske, P. Norberg, I. K. Baldry, S. P. Bamford, A. M. Hopkins, J. Loveday, J. A. Peacock, E. Andrae, J. Bland-Hawthorn, S. Brough, M. J. I. Brown, E. Cameron, J. H. Y. Ching, M. Colless, C. J. Conselice, S. M. Croom, N. J. G. Cross, R. de Propris, S. Dye, M. J. Drinkwater, S. Ellis, A. W. Graham, M. W. Grootes, M. Gunawardhana, D. H. Jones, E. van Kampen, C. Maraston, R. C. Nichol, H. R. Parkinson, S. Phillipps, K. Pimbblet, C. C. Popescu, M. Prescott, I. G. Roseboom, E. M. Sadler, A. E. Sansom, R. G. Sharp, D. J. B. Smith, E. Taylor, D. Thomas, R. J. Tuffs, D. Wijesinghe, L. Dunne, C. S. Frenk, M. J. Jarvis, B. F. Madore, M. J. Meyer, M. Seibert, L. Staveley-Smith, W. J. Sutherland, and S. J. Warren. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413:971–995, May 2011. doi: 10.1111/j.1365-2966.2010.18188.x.

[80] M. Colless, G. Dalton, S. Maddox, W. Sutherland, P. Norberg, S. Cole, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, K. Deeley, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, D. Madgwick, J. A. Peacock, B. A. Peterson, I. Price, M. Seaborne, and K. Taylor. The 2dF Galaxy Redshift Survey: spectra and redshifts. *MNRAS*, 328:1039–1063, December 2001. doi: 10.1046/j.1365-8711.2001.04902.x.

[81] S. Eales, L. Dunne, D. Clements, A. Cooray, G. de Zotti, S. Dye, R. Ivison, M. Jarvis, G. Lagache, S. Maddox, M. Negrello, S. Serjeant, M. A. Thompson, E. van Kampen, A. Amblard, P. Andreani, M. Baes, A. Beelen, G. J. Bendo, D. Benford, F. Bertoldi, J. Bock, D. Bonfield, A. Boselli, C. Bridge, V. Buat, D. Burgarella, R. Carlberg, A. Cava, P. Chanial, S. Charlot, N. Christopher, P. Coles, L. Cortese, A. Dariush, E. da Cunha, G. Dalton, L. Danese, H. Dannerbauer, S. Driver, J. Dunlop, L. Fan, D. Farrah, D. Frayer, C. Frenk, J. Geach, J. Gardner, H. Gomez, J. González-Nuevo, E. González-Solares, M. Griffin, M. Hardcastle, E. Hatziminaoglou, D. Herranz, D. Hughes, E. Ibar, W.-S. Jeong, C. Lacey, A. Lapi, A. Lawrence, M. Lee, L. Leeuw, J. Liske, M. López-Caniego, T. Müller,

K. Nandra, P. Panuzzo, A. Papageorgiou, G. Patanchon, J. Peacock, C. Pearson, S. Phillipps, M. Pohlen, C. Popescu, S. Rawlings, E. Rigby, M. Rigopoulou, A. Robotham, G. Rodighiero, A. Sansom, B. Schulz, D. Scott, D. J. B. Smith, B. Sibthorpe, I. Smail, J. Stevens, W. Sutherland, T. Takeuchi, J. Tedds, P. Temi, R. Tuffs, M. Trichas, M. Vaccari, I. Valtchanov, P. van der Werf, A. Verma, J. Vieria, C. Vlahakis, and G. J. White. The Herschel ATLAS. *PASP*, 122:499–515, May 2010. doi: 10.1086/653086.

[82] N. C. Hambly, R. S. Collins, N. J. G. Cross, R. G. Mann, M. A. Read, E. T. W. Sutorius, I. Bond, J. Bryant, J. P. Emerson, A. Lawrence, L. Rimoldini, J. M. Stewart, P. M. Williams, A. Adamson, P. Hirst, S. Dye, and S. J. Warren. The WFCAM Science Archive. *MNRAS*, 384:637–662, February 2008. doi: 10.1111/j.1365-2966.2007.12700.x.

[83] J. P. Emerson, M. J. Irwin, J. Lewis, S. Hodgkin, D. Evans, P. Bunclark, R. McMahon, N. C. Hambly, R. G. Mann, I. Bond, E. Sutorius, M. Read, P. Williams, A. Lawrence, and M. Stewart. VISTA data flow system: overview. In P. J. Quinn & A. Bridger, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5493 of *The Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 401–410, September 2004. doi: 10.1117/12.551582.

[84] M. J. Irwin, J. Lewis, S. Hodgkin, P. Bunclark, D. Evans, R. McMahon, J. P. Emerson, M. Stewart, and S. Beard. VISTA data flow system: pipeline processing for WFCAM and VISTA. In P. J. Quinn & A. Bridger, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5493 of *The Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 411–422, September 2004. doi: 10.1117/12.551449.

[85] S. T. Hodgkin, M. J. Irwin, P. C. Hewett, and S. J. Warren. The UKIRT wide field camera ZYJHK photometric system: calibration from 2MASS. *MNRAS*, 394:675–692, April 2009. doi: 10.1111/j.1365-2966.2008.14387.x.

[86] P. J. Francis, P. C. Hewett, C. B. Foltz, F. H. Chaffee, R. J. Weymann, and S. L. Morris. A high signal-to-noise ratio composite quasar spectrum. *ApJ*, 373:465–470, June 1991. doi: 10.1086/170066.

[87] J. J. Bochanski, S. L. Hawley, K. R. Covey, A. A. West, I. N. Reid, D. A. Golimowski, and Ž. Ivezić. The Luminosity and Mass Functions of Low-mass Stars in the Galactic Disk. II. The Field. *AJ*, 139:2679–2699, June 2010. doi: 10.1088/0004-6256/139/6/2679.

[88] N. Lodieu, B. Burningham, N. C. Hambly, and D. J. Pinfield. Identifying nearby field T dwarfs in the UKIDSS Galactic Clusters Survey. *MNRAS*, 397:258–264, July 2009. doi: 10.1111/j.1365-2966.2008.14384.x.

[89] B. Burningham, D. J. Pinfield, P. W. Lucas, S. K. Leggett, N. R. Deacon, M. Tamura, C. G. Tinney, N. Lodieu, Z. H. Zhang, N. Huelamo, H. R. A. Jones, D. N. Murray, D. J. Mortlock, M. Patel, D. Barrado Y Navascués, M. R. Zapatero Osorio, M. Ishii, M. Kuzuhara, and R. L. Smart. 47 new T dwarfs from the UKIDSS Large Area Survey. *MNRAS*, 406:1885–1906, August 2010. doi: 10.1111/j.1365-2966.2010.16800.x.

[90] X. Fan. Simulation of Stellar Objects in SDSS Color Space. *AJ*, 117:2528–2551, May 1999. doi: 10.1086/300848.

[91] R. H. Lupton, J. E. Gunn, and A. S. Szalay. A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements. *AJ*, 118:1406–1410, September 1999. doi: 10.1086/301004.

[92] D. M. Peterson, C. A. Hummel, T. A. Pauls, J. T. Armstrong, J. A. Benson, G. C. Gilbreath, R. B. Hindsley, D. J. Hutter, K. J. Johnston, D. Mozurkewich, and H. R. Schmitt. Vega is a rapidly rotating star. *Nature*, 440:896–899, April 2006. doi: 10.1038/nature04661.

[93] J. Maíz-Apellániz. A Uniform Set of Optical/NIR Photometric Zero Points to be Used with CHORIZOS. In C. Sterken, editor, *The Future of Photometric, Spectrophotometric and Polarimetric Standardization*, volume 364 of *Astronomical Society of the Pacific Conference Series*, page 227, April 2007.

[94] A. T. Tokunaga and W. D. Vacca. The Mauna Kea Observatories Near-Infrared Filter Set. III. Isophotal Wavelengths and Absolute Calibration. *PASP*, 117:421–426, April 2005. doi: 10.1086/429382.

[95] M. Cohen, R. G. Walker, M. J. Barlow, and J. R. Deacon. Spectral irradiance calibration in the infrared. I - Ground-based and IRAS broadband calibrations. *AJ*, 104:1650–1657, October 1992. doi: 10.1086/116349.

[96] R. C. Bohlin and R. L. Gilliland. Hubble Space Telescope Absolute Spectrophotometry of Vega from the Far-Ultraviolet to the Infrared. *AJ*, 127:3508–3515, June 2004. doi: 10.1086/420715.

[97] R. C. Bohlin. HST Stellar Standards with 1% Accuracy in Absolute Flux. In C. Sterken, editor, *The Future of Photometric, Spectrophotometric and Polarimetric Standardization*, volume 364 of *Astronomical Society of the Pacific Conference Series*, page 315, April 2007.

[98] W. L. Wiese, M. W. Smith, and B. M. Glennon. *Atomic transition probabilities. Vol.: Hydrogen through Neon. A critical data compilation*. US Department of Commerce, National Buereau of Standards, 1966, NSRDS-NBS 4, Washington, D.C., 1966.

[99] D. P. Schneider, P. B. Hall, G. T. Richards, et al. The Sloan Digital Sky Survey Quasar Catalog. IV. Fifth Data Release. *AJ*, 134:102–117, July 2007. doi: 10.1086/518474.

[100] G. T. Richards, X. Fan, H. J. Newberg, M. A. Strauss, D. E. Vanden Berk, D. P. Schneider, B. Yanny, A. Boucher, S. Burles, J. A. Frieman, J. E. Gunn, P. B. Hall, Ž. Ivezić, S. Kent, J. Loveday, R. H. Lupton, C. M. Rockosi, D. J. Schlegel, C. Stoughton, M. SubbaRao, and D. G. York. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Quasar Sample. *AJ*, 123:2945–2975, June 2002. doi: 10.1086/340187.

[101] X. Fan, V. K. Narayanan, R. H. Lupton, M. A. Strauss, G. R. Knapp, R. H. Becker, R. L. White, L. Pentericci, S. K. Leggett, Z. Haiman, J. E. Gunn, Ž. Ivezić, D. P. Schneider, S. F. Anderson, J. Brinkmann, N. A. Bahcall, A. J. Connolly, I. Csabai, M. Doi, M. Fukugita, T. Geballe, E. K. Grebel, D. Harbeck, G. Hennessy, D. Q. Lamb, G. Miknaitis, J. A. Munn, R. Nichol, S. Okamura, J. R. Pier, F. Prada, G. T. Richards, A. Szalay, and D. G. York. A Survey of $z > 5.8$ Quasars in the Sloan Digital Sky Survey. I. Discovery of Three New Quasars and the Spatial Density of Luminous Quasars at $z \sim 6$. *AJ*, 122:2833–2849, December 2001. doi: 10.1086/324111.

[102] D. J. Mortlock, M. Patel, S. J. Warren, B. P. Venemans, R. G. McMahon, P. Hewett, C. Simpson, and R. G. Sharp. The UKIRT Infrared Deep Sky Survey and the search for the most distant quasars. *ArXiv e-prints*, October 2008.

[103] D. J. Pinfield, B. Burningham, M. Tamura, et al. Fifteen new T dwarfs discovered in the UKIDSS Large Area Survey. *MNRAS*, 390:304–322, October 2008. doi: 10.1111/j.1365-2966.2008.13729.x.

[104] M. Rowan-Robinson, T. Babbedge, S. Oliver, et al. Photometric redshifts in the SWIRE Survey. *MNRAS*, 386:697–714, May 2008. doi: 10.1111/j.1365-2966.2008.13109.x.

[105] J. D. Kirkpatrick. New Spectral Types L and T. *ARA&A* , 43:195–245, September 2005. doi: 10.1146/annurev.astro.42.053102.134017.

[106] E. Glikman, S. G. Djorgovski, D. Stern, A. Dey, B. T. Jannuzi, and K.-S. Lee. The Faint End of the Quasar Luminosity Function at $z \sim 4$: Implications for Ionization of the Intergalactic Medium and Cosmic Downsizing. *ApJL*, 728:L26, February 2011. doi: 10.1088/2041-8205/728/2/L26.

[107] B. J. Boyle, T. Shanks, S. M. Croom, R. J. Smith, L. Miller, N. Loaring, and C. Heymans. The 2dF QSO Redshift Survey - I. The optical luminosity function of quasi-stellar objects. *MNRAS*, 317:1014–1022, October 2000. doi: 10.1046/j.1365-8711.2000.03730.x.

[108] G. T. Richards, M. A. Strauss, X. Fan, et al. The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3. *AJ*, 131:2766–2787, June 2006. doi: 10.1086/503559.

[109] M. Schmidt, D. P. Schneider, and J. E. Gunn. Spectrscopic CCD Surveys for Quasars at Large Redshift. IV. Evolution of the Luminosity Function from Quasars Detected by Their Lyman-Alpha Emission. *AJ*, 110:68, July 1995. doi: 10.1086/117497.

[110] S. M. Croom, G. T. Richards, T. Shanks, B. J. Boyle, M. A. Strauss, A. D. Myers, R. C. Nichol, K. A. Pimbblet, N. P. Ross, D. P. Schneider, R. G. Sharp, and D. A. Wake. The 2dF-SDSS LRG and QSO survey: the QSO luminosity function at $0.4 < z < 2.6$. *MNRAS*, 399:1755–1772, November 2009. doi: 10.1111/j.1365-2966.2009.15398.x.

[111] A. J. Barger, L. L. Cowie, R. F. Mushotzky, Y. Yang, W.-H. Wang, A. T. Steffen, and P. Capak. The Cosmic Evolution of Hard X-Ray-selected Active Galactic Nuclei. *AJ*, 129:578–609, February 2005. doi: 10.1086/426915.

[112] L. L. Cowie, A. Songaila, E. M. Hu, and J. G. Cohen. New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields. *AJ*, 112:839, September 1996. doi: 10.1086/118058.

[113] Y. Li, L. Hernquist, B. Robertson, T. J. Cox, P. F. Hopkins, V. Springel, L. Gao, T. Di Matteo, A. R. Zentner, A. Jenkins, and N. Yoshida. Formation of $z \sim 6$ Quasars from Hierarchical Galaxy Mergers. *ApJ*, 665:187–208, August 2007. doi: 10.1086/519297.

[114] D. Sijacki, V. Springel, and M. G. Haehnelt. Growing the first bright quasars in cosmological simulations of structure formation. *MNRAS*, 400:100–122, November 2009. doi: 10.1111/j.1365-2966.2009.15452.x.

[115] S. M. Croom, R. J. Smith, B. J. Boyle, T. Shanks, L. Miller, P. J. Outram, and N. S. Loaring. The 2dF QSO Redshift Survey - XII. The spectroscopic catalogue and luminosity function. *MNRAS*, 349:1397–1418, April 2004. doi: 10.1111/j.1365-2966.2004.07619.x.

[116] M. P. Hunt, C. C. Steidel, K. L. Adelberger, and A. E. Shapley. The Faint End of the QSO Luminosity Function at z = 3. *ApJ*, 605:625–630, April 2004. doi: 10.1086/381727.

[117] A. C. Robin, C. Reylé, S. Derrière, and S. Picaud. A synthetic view on structure and evolution of the Milky Way. *A&A*, 409:523–540, October 2003. doi: 10.1051/0004-6361:20031117.

[118] A. C. Robin and C. Reylé. The Initial Mass Function at Low Masses. In J. M. De Buizer & N. S. van der Bliek, editor, *Galactic Star Formation Across the Stellar Mass Spectrum*, volume 287 of *Astronomical Society of the Pacific Conference Series*, pages 104–109, 2003.

[119] M. Schultheis, A. C. Robin, C. Reylé, H. J. McCracken, E. Bertin, Y. Mellier, and O. Le Fèvre. Stellar populations in the CFHTLS. I. New constraints on the IMF at low mass. *A&A*, 447:185–198, February 2006. doi: 10.1051/0004-6361:20053376.

[120] M. Haywood, A. C. Robin, and M. Creze. The evolution of the Milky Way disc. II. Constraints from star counts at the galactic poles. *A&A*, 320:440–459, April 1997.

[121] T. Lejeune, F. Cuisinier, and R. Buser. Standard stellar library for evolutionary synthesis. I. Calibration of theoretical spectra. *A&AS*, 125:229–246, October 1997. doi: 10.1051/aas:1997373.

[122] F. Allard, P. H. Hauschildt, D. R. Alexander, and S. Starrfield. Model Atmospheres of Very Low Mass Stars and Brown Dwarfs. *ARA&A* , 35:137–177, 1997. doi: 10.1146/annurev.astro.35.1.137.

[123] J. N. Bahcall and R. M. Soneira. Predicted star counts in selected fields and photometric bands Applications to galactic structure, the disk luminosity function, and the detection of a massive halo. *ApJS*, 47:357–403, December 1981. doi: 10.1086/190765.

[124] G. Gilmore. New light on faint stars. VI - Structure and evolution of the Galactic spheroid. *MNRAS*, 207:223–240, March 1984.

[125] M. Patel. *The search for High Redshift Quasars in UKIDSS*. PhD thesis, Imperial College London, 2010.

[126] N. Maddox, P. C. Hewett, S. J. Warren, and S. M. Croom. Luminous K-band selected quasars from UKIDSS. *MNRAS*, 386:1605–1624, May 2008. doi: 10.1111/j.1365-2966.2008.13138.x.

[127] G. T. Richards, R. C. Nichol, A. G. Gray, R. J. Brunner, R. H. Lupton, D. E. Vanden Berk, S. S. Chong, M. A. Weinstein, D. P. Schneider, S. F. Anderson, J. A. Munn, H. C. Harris, M. A. Strauss, X. Fan, J. E. Gunn, Ž. Ivezić, D. G. York, J. Brinkmann, and A. W. Moore. Efficient Photometric Selection of Quasars from the Sloan Digital Sky Survey: 100,000 z > 3 Quasars from Data Release One. *ApJS*, 155:257–269, December 2004. doi: 10.1086/425356.

[128] G. T. Richards, R. P. Deo, M. Lacy, A. D. Myers, R. C. Nichol, N. L. Zakamska, R. J. Brunner, W. N. Brandt, A. G. Gray, J. K. Parejko, A. Ptak, D. P. Schneider, L. J. Storrie-Lombardi, and A. S. Szalay. Eight-Dimensional Mid-Infrared/Optical

Bayesian Quasar Selection. *AJ*, 137:3884–3899, April 2009. doi: 10.1088/0004-6256/137/4/3884.

[129] J. Bovy, J. F. Hennawi, D. W. Hogg, A. D. Myers, J. A. Kirkpatrick, D. J. Schlegel, N. P. Ross, E. S. Sheldon, I. D. McGreer, D. P. Schneider, and B. A. Weaver. Think Outside the Color Box: Probabilistic Target Selection and the SDSS-XDQSO Quasar Targeting Catalog. *ApJ*, 729:141, March 2011. doi: 10.1088/0004-637X/729/2/141.

[130] D. J. Mortlock, M. Patel, S. J. Warren, P. C. Hewett, B. P. Venemans, R. G. McMahon, and C. Simpson. Probabilistic selection of high-redshift quasars. *MNRAS*, 419:390–410, January 2012. doi: 10.1111/j.1365-2966.2011.19710.x.

[131] M. Irwin, M. Riello, S. Hodgkin, E. Gonzales-Solares, D. W. Evans, and P. Bunclark. Pipeline processing of wide-field near-infrared data from WFCAM. *MNRAS*, in preparation, .

[132] Z. Ivezic, J. A. Tyson, T. Axelrod, D. Burke, C. F. Claver, K. H. Cook, S. M. Kahn, R. H. Lupton, D. G. Monet, P. A. Pinto, M. A. Strauss, C. W. Stubbs, L. Jones, A. Saha, R. Scranton, C. Smith, and LSST Collaboration. LSST: From Science Drivers To Reference Design and Anticipated Data Products. In *American Astronomical Society Meeting Abstracts No.213*, volume 41 of *Bulletin of the American Astronomical Society*, page 460.03, January 2009.

[133] R. J. Laureijs, L. Duvet, I. Escudero Sanz, P. Gondoin, D. H. Lumb, T. Oosterbroek, and G. Saavedra Criado. The Euclid Mission. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7731 of *The Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, July 2010. doi: 10.1117/12.857123.

[134] C. Yèche, P. Petitjean, J. Rich, E. Aubourg, N. Busca, J.-C. Hamilton, J.-M. Le Goff, I. Paris, S. Peirani, C. Pichon, E. Rollinde, and M. Vargas-Magaña. Artificial neural networks for quasar selection and photometric redshift determination. *A&A*, 523:A14, November 2010. doi: 10.1051/0004-6361/200913508.

[135] X. Fan, M. A. Strauss, D. P. Schneider, R. H. Becker, R. L. White, Z. Haiman, M. Gregg, L. Pentericci, E. K. Grebel, V. K. Narayanan, Y.-S. Loh, G. T. Richards, J. E. Gunn, R. H. Lupton, G. R. Knapp, Ž. Ivezić, W. N. Brandt, M. Collinge, L. Hao, D. Harbeck, F. Prada, J. Schaye, I. Strateva, N. Zakamska, S. Anderson, J. Brinkmann, N. A. Bahcall, D. Q. Lamb, S. Okamura, A. Szalay, and D. G. York. A Survey of $z > 5.7$ Quasars in the Sloan Digital Sky Survey. II. Discovery of Three Additional Quasars at $z > 6$. *AJ*, 125:1649–1659, April 2003. doi: 10.1086/368246.

[136] L. Jiang, X. Fan, J. Annis, R. H. Becker, R. L. White, K. Chiu, H. Lin, R. H. Lupton, G. T. Richards, M. A. Strauss, S. Jester, and D. P. Schneider. A Survey of z $\sim$ 6 Quasars in the Sloan Digital Sky Survey Deep Stripe. I. A Flux-Limited Sample at $z_{AB} < 21$. *AJ*, 135:1057–1066, March 2008. doi: 10.1088/0004-6256/135/3/1057.

[137] B. Buzzoni, B. Delabre, H. Dekker, S. Dodorico, D. Enard, P. Focardi, B. Gustafsson, W. Nees, J. Paureau, and R. Reiss. The ESO Faint Object Spectrograph and Camera (EFOSC). *The Messenger*, 38:9–13, December 1984.

[138] European Southern Observatory. ESO Exposure Time Calculators & Simulations, . URL `http://www.eso.org/observing/etc/`.

[139] J. F. Kenney and E. S. Keeping. *Relation Between Mean, Median, and Mode*, volume 1, pages 32–35, 52–54, 211–212. Van Nostrand, Princton: NJ, 3rd edition, 1962.

[140] European Southern Observatory. Spectroscopic Standard Stars and Imaging Standards for Gunn & interference filters, . URL `http://www.eso.org/sci/facilities/paranal/instruments/fors/tools/FORS_Std/FORS_Std.html`.

[141] F. Patat, S. Moehler, K. O'Brien, E. Pompei, T. Bensby, G. Carraro, A. de Ugarte Postigo, A. Fox, I. Gavignaud, G. James, H. Korhonen, C. Ledoux, S. Randall, H. Sana, J. Smoker, S. Stefl, and T. Szeifert. Optical atmospheric extinction over Cerro Paranal. *A&A*, 527:A91, March 2011. doi: 10.1051/0004-6361/201015537.

[142] Monaco, L. and Snodgrass, C. *EFOSC2 User's Manual Vol. 3.1*. European Southern Observatory, 3.1 edition, 2008.

[143] H. Tüg. Vertical Extinction on La Silla. *The Messenger*, 11:7–8, December 1977.

[144] P. B. Stetson, L. E. Davis, and D. R. Crabtree. Future development of the DAOPHOT crowded-field photometry package. In G. H. Jacoby, editor, *CCDs in astronomy*, volume 8 of *Astronomical Society of the Pacific Conference Series*, pages 289–304, 1990.

[145] W. J. Merline and S. B. Howell. A Realistic Model for Point-sources Imaged on Array Detectors: The Model and Initial Results. *Experimental Astronomy*, 6:163–210, January 1995. doi: 10.1007/BF00421131.

[146] S. B. Howell. *Handbook of CCD astronomy*. Cambridge University Press, Cambridge, UK, 2006.

[147] E. Davis, L. *A user's guide to IRAF Apphot package*. National Optical Astronomy Observatories, 1989.

[148] N. Epchtein, B. de Batz, L. Capoani, L. Chevallier, E. Copet, P. Fouqué, P. Lacombe, T. Le Bertre, S. Pau, D. Rouan, S. Ruphy, G. Simon, D. Tiphène, W. B. Burton, E. Bertin, E. Deul, H. Habing, J. Borsenberger, M. Dennefeld, F. Guglielmo, C. Loup, G. Mamon, Y. Ng, A. Omont, L. Provost, J.-C. Renault, F. Tanguy, S. Kimeswenger, C. Kienel, F. Garzon, P. Persi, M. Ferrari-Toniolo, A. Robin, G. Paturel, I. Vauglin, T. Forveille, X. Delfosse, J. Hron, M. Schultheis, I. Appenzeller, S. Wagner, L. Balazs, A. Holl, J. Lépine, P. Boscolo, E. Picazzio, P.-A. Duc, and M.-O. Mennessier. The deep near-infrared southern sky survey (DENIS). *The Messenger*, 87:27–34, March 1997.

[149] I. Appenzeller, K. Fricke, W. Fürtig, W. Gässler, R. Häfner, R. Harke, H.-J. Hess, W. Hummel, P. Jürgens, R.-P. Kudritzki, K.-H. Mantel, W. Meisl, B. Muschielok, H. Nicklas, G. Rupprecht, W. Seifert, O. Stahl, T. Szeifert, and K. Tarantik. Successful commissioning of FORS1 - the first optical instrument on the VLT. *The Messenger*, 94:1–6, December 1998.

[150] B. P. Venemans, J. R. Findlay, R. G. Mcmahon, and W. J. Sutherland. *MNRAS*, in preparation, 2011.

[151] J. Neyman. Outline of a theory of statistical estimation based on the classical theory of probability. *Phil. Trans. R. Soc. Lond. A* , 236:333–380, 1937.

[152] J. Miralda-Escude. Reionization of the Intergalactic Medium and the Damping Wing of the Gunn-Peterson Trough. *ApJ*, 501:15, July 1998. doi: 10.1086/305799.

[153] J. S. Bolton, M. G. Haehnelt, S. J. Warren, P. C. Hewett, D. J. Mortlock, B. P. Venemans, R. G. McMahon, and C. Simpson. How neutral is the intergalactic medium surrounding the redshift z = 7.085 quasar ULAS J1120+0641? *MNRAS*, 416:L70–L74, September 2011. doi: 10.1111/j.1745-3933.2011.01100.x.

[154] S. Cantalupo, C. Porciani, and S. J. Lilly. Mapping Neutral Hydrogen during Reionization with the Lyα Emission from Quasar Ionization Fronts. *ApJ*, 672: 48–58, January 2008. doi: 10.1086/523298.

[155] J. A Nelder and R. Mead. A Simplex Method for Function Minimization,. *Comp. Journ.*, 7:308–313, 1965.